Protocol

This trial protocol has been provided by the authors to give readers additional information about their work.

Protocol for: Choueiri TK, Powles T, Burotto M, et al. Nivolumab plus cabozantinib versus sunitinib for advanced renal-cell carcinoma. N Engl J Med 2021;384:829-41. DOI: 10.1056/NEJMoa2026982

This supplement contains the following items:

1. Original CheckMate 9ER protocol (page 2), final protocol (page 160), summary of changes (page 162)
2. Original CheckMate 9ER statistical analysis plan (page 332), final statistical analysis plan (page 402), summary of changes (page 463).

Protocol Number: CA2099ER
IND Number: 134,374
EUDRACT Number 2017-000759-20
Date 08-Mar-2017

Clinical Protocol CA2099ER

A Phase 3, Randomized, Open-Label Study of Nivolumab Combined with Cabozantinib or Nivolumab and Ipilimumab Combined with Cabozantinib versus Sunitinib in Participants with Previously Untreated, Advanced or Metastatic Renal Cell Carcinoma

(CheckMate 9ER: CHECKpoint pathway and nivoluMAb clinical Trial Evaluation 9ER)

Study Director/ Medical Monitor

24-hr Emergency Telephone Number
USA: 1-866-470-2267
International: +1-248-844-7390

Bristol-Myers Squibb Research and Development
3401 Princeton Pike
Lawrence Township, NJ 08648
Avenue de Finlande 4
B-1420 Braine-l'Alleud, Belgium

DOCUMENT HISTORY

Document	Date of Issue	Summary of Change
Original Protocol	08-Mar-2017	Not applicable

TABLE OF CONTENTS

1 SYNOPSIS

Protocol Title: A Phase 3, Randomized, Open-Label Study of Nivolumab Combined with Cabozantinib or Nivolumab and Ipilimumab Combined with Cabozantinib versus Sunitinib in Participants with Previously Untreated, Advanced or Metastatic Renal Cell Carcinoma

Study phase: 3

Rationale: Although multiple agents are approved as monotherapies for the treatment of patients with metastatic renal cell carcinoma (mRCC), the testing of combination therapies, in particular, treatment with immune-checkpoint inhibitors with tyrosine kinase inhibitors (TKIs) has not been fully explored. While single agent therapies have improved outcomes, ongoing drug resistance and disease progression demonstrate an urgent need to find more effective therapies for mRCC patients. The top priorities in treating RCC continue to be improving progression free survival (PFS) and overall survival (OS), management of toxicities, and a better understanding of biomarkers. The CA2099ER trial will include previously untreated participants with mRCC and uses 2 well-characterized immune checkpoint inhibitors (nivolumab and ipilimumab) in combination with cabozantinib, a known standard-of-care in previously treated mRCC participants. Nivolumab combined with cabozantinib or nivolumab and ipilimumab combined with cabozantinib may be an important step forward in evaluating combination regimens which could potentially optimize the management of previously untreated participants with mRCC.

Study Population: Male and female participants ≥ 18 years or the age of majority with previously untreated, advanced or metastatic renal cell carcinoma (RCC).

The following list contains key eligibility criteria only. For full list of eligibility criteria please see Section 6.

Key Inclusion Criteria

- Histological confirmation of RCC with a clear-cell component, including participants who may also have sarcomatoid features
- Advanced (not amenable to curative surgery or radiation therapy) or metastatic (American Joint Committee on Cancer [AJCC] Stage IV) RCC
- No prior systemic therapy for RCC with the following exception:
    - One prior adjuvant or neoadjuvant therapy for completely resectable RCC if such therapy did not include an agent that targets vascular endothelial growth factor (VEGF) or VEGF receptors and if recurrence occurred at least 6 months after the last dose of adjuvant or neoadjuvant therapy
- Karnofsky Performance Status (KPS) grade ≥ 70%
- Measurable disease as per RECIST v1.1 per investigator
- Tumor tissue, preferably obtained within 3 months but no more than 12 months prior to enrollment, with an associated pathology report, must be received by the central laboratory during screening for determination of PD-L1 expression. In order to be randomized, a participant must be classified as PD-L1 expression ≥ 1%, PD-L1 expression < 1%, or PD-L1 expression indeterminate.

- Participants with favorable, intermediate, and poor risk categories will be eligible for the study. Participants must be categorized according to favorable versus intermediate versus poor risk status at registration as per International Metastatic RCC Database Consortium (IMDC) criteria.
- Negative pregnancy test and able to meet protocol-specified reproductive requirements

Key Exclusion Criteria

- Any active central nervous system (CNS) metastases. Participants with treated, stable CNS metastases for at least 3 months are eligible
- Any tumor invading the superior vena cava (SVC), other major blood vessels, or GI tract; any evidence of endotracheal or endobronchial tumor
- Prior systemic treatment with VEGF, MET, AXL, KIT or RET targeted therapy (including, but not limited to, sunitinib, pazopanib, axitinib, tivozanib, sorafenib, lenvatinib, bevacizumab, and cabozantinib)
- Prior treatment with an anti-PD-1, anti-PD-L1, anti-PD-L2, anti-CD137, or anti-CTLA-4 antibody, or any other antibody or drug specifically targeting T-cell co-stimulation or checkpoint pathways
- Any active, known or suspected autoimmune disease or any condition requiring systemic treatment with corticosteroids (> 10 mg daily prednisone equivalent) or other immunosuppressive medications within 14 days of randomization
- Uncontrolled adrenal insufficiency
- Poorly controlled hypertension despite antihypertensive therapy
- History of unstable angina, myocardial infarction, symptomatic peripheral vascular disease, congestive heart failure (CHF, Class III or IV as defined by the New York Heart Association [NYHA]), or cerebrovascular accident (CVA)
- Deep vein thrombosis (DVT) or pulmonary embolism (PE) unless adequately treated with low molecular weight heparin (LMWH)
- Any unstable cardiac arrhythmia; prolonged QTcF > 450 msec for males and > 470 msec for females
- Serious, non-healing wound or ulcer; evidence of active bleeding or bleeding susceptibility; history of abdominal fistula, gastrointestinal perforation, intra-abdominal abscess, bowel or gastric outlet obstruction
- Concomitant strong CYP3A4 inducers or inhibitors within 14 days prior to randomization
- Ejection fraction ≤ 50% on screening echocardiogram or multigated acquisition scan (MUGA)
- Major surgery (eg, nephrectomy) less than 6 weeks prior to randomization

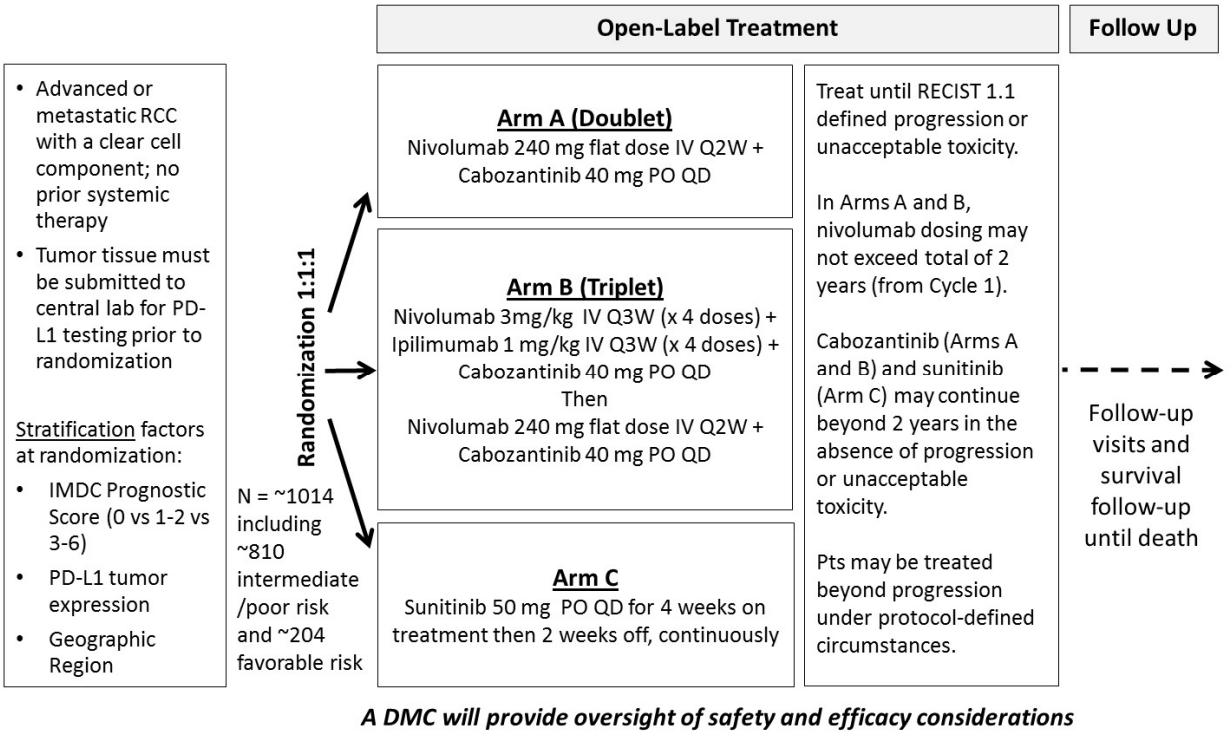
Objectives and Endpoints:

Objective	Endpoint
Primary	
To compare progression-free survival (PFS) per blinded independent central review (BICR) of Arm A (doublet) with Arm C (sunitinib) and of Arm B (triplet) with Arm C (sunitinib) in all intermediate/poor risk randomized participants.	The primary objective specifies two comparisons: PFS per BICR of Arm A versus Arm C and of Arm B versus Arm C in all intermediate/poor risk randomized participants. PFS is defined as the time between the date of randomization and the first date of the documented progression, or death due to any cause, whichever occurs first. Participants who die without a reported progression will be considered to have progressed on the date of their death. Participants who did not progress or die will be censored on the date of their last evaluable tumor assessment. Participants who did not have any on study tumor assessments and did not die will be censored on their date of randomization. Participants who started anti-cancer therapy without a prior reported progression will be censored on the date of their last evaluable tumor assessment prior to the initiation of first subsequent anti-cancer therapy.
Secondary	
To compare progression-free survival (PFS) per BICR of Arm A (doublet) with Arm C (sunitinib) and of Arm B (triplet) with Arm C (sunitinib) in all randomized participants.	This secondary objective specifies two comparisons: PFS per BICR of Arm A versus Arm C and of Arm B versus Arm C in all randomized participants. The PFS will be defined similarly to the primary endpoint except that all randomized participants (any risk participants) will be used.
To compare overall survival (OS) of Arm A with Arm C and of Arm B with Arm C in all intermediate/poor risk randomized participants.	The secondary endpoint specifies two comparisons: OS of Arm A versus Arm C and of Arm B versus Arm C in all intermediate/poor risk randomized participants. OS is defined as the time between the date of randomization and the date of death due to any cause. A participant who has not died will be censored at the last known alive date.
To compare overall survival (OS) of Arm A with Arm C and of Arm B with Arm C in all randomized participants.	This secondary objective specifies two comparisons: OS of Arm A versus Arm C and of Arm B versus Arm C in all randomized participants. The OS will be defined similarly to the first secondary endpoint except that all randomized participants (any risk participants) will be used.
To evaluate the objective response rate (ORR) per BICR in all intermediate/poor risk randomized and all randomized participants.	This secondary objective is ORR per BICR in all intermediate/poor risk randomized participants and in all randomized participants (any risk participants). ORR is defined as the proportion of randomized participants who achieve a best response of complete response (CR) or partial response (PR) using the RECIST 1.1 criteria. Best overall response (BOR) is defined as the best response designation recorded between the date of randomization and the date of objectively documented progression per RECIST 1.1 or the date of subsequent therapy (including tumor-directed radiotherapy and tumor-directed surgery), whichever occurs first. For participants without documented progression or subsequent therapy, all available response designations will contribute to the BOR assessment. Duration of response (DOR) is defined as the time between the date of first confirmed documented response (CR or PR) to the date of first documented tumor progression (per RECIST 1.1) or death due to any cause, whichever occurs first. Participants

Objective	Endpoint
	who neither progress nor die will be censored on the date of their last tumor assessment. Responders who started anti-cancer therapy without a prior reported progression will be censored on the date of their last evaluable tumor assessment prior to the initiation of first subsequent anti-cancer therapy. Time to response (TTR) is defined as the time from randomization to the date of the first confirmed documented response (CR or PR), as assessed by BICR. DOR and TTR will be evaluated for responders (CR or PR) only.
To assess overall safety and tolerability in all treated participants.	As measured by the incidence of adverse events (AEs), serious adverse events (SAEs), AEs leading to discontinuation, deaths, laboratory abnormalities and changes from baseline.

Overall Design: This is an open label, randomized trial of nivolumab combined with cabozantinib (doublet regimen) or nivolumab + ipilimumab combined with cabozantinib (triplet regimen) versus sunitinib in participants with previously untreated (first line) advanced or metastatic RCC. Participants will be randomized between 3 arms in a 1:1:1 ratio with approximately 810 intermediate/poor risk participants (270 per arm) and 204 favorable risk participants (68 per arm). Participants will be stratified for randomization by IMDC prognostic score (0 [favorable risk] versus 1-2 [intermediate risk] versus 3-6 [poor risk]), PD-L1 tumor expression (≥ 1% versus < 1% or indeterminate), and region (US/Canada/Western Europe/Northern Europe versus rest of the world [ROW]).

The study design schematic is presented below.



Abbreviations: BICR= blinded independent central review; DMC= data monitoring committee; IMDC= International Metastatic Renal Cell Carcinoma Database Consortium; IV=intravenous; PD= progressive disease; PD-L1=

programmed death-ligand 1; PO= orally by mouth; Pts=participants; Q2W=every 2 weeks; Q3W=every 3 weeks; QD= once daily; RCC=renal cell carcinoma.

Number of Participants: Approximately 1353 participants will enroll in order to randomize 1014 participants (338 per arm). This includes approximately 204 favorable risk randomized (68 per arm) and approximately 810 intermediate/poor risk randomized participants (270 per each arm). The number of randomized participants with favorable risk disease will be capped at approximately 204 participants.

Treatment Arms and Duration:

- Arm A (Doublet): Nivolumab 240 mg flat dose intravenously (IV) every 2 weeks (Q2W) + Cabozantinib 40 mg orally by mouth (PO) once daily (QD)
    - Nivolumab treatment until disease progression or unacceptable toxicity with maximum treatment of 2 years
    - Cabozantinib treatment until disease progression or unacceptable toxicity
- Arm B (Triplet): Nivolumab 3mg/kg IV Q3W x 4 doses + Ipilimumab 1 mg/kg IV Q3W x 4 doses + Cabozantinib 40 mg PO QD
    - Then, Nivolumab 240 mg flat dose IV Q2W + Cabozantinib 40 mg PO QD
    - Nivolumab to be continued until disease progression or unacceptable toxicity with maximum treatment of 2 years from the start of first dose in Cycle 1
    - Cabozantinib until disease progression or unacceptable toxicity
- Arm C: Sunitinib 50 mg PO QD for 4 weeks, followed by 2 weeks off-treatment, per cycle. Cycles to be continued until progression or unacceptable toxicity

Refer to Section 7.1 Treatments Administered for additional details.

Study treatment:

Study Drugs for CA2099ER

Medication	Potency	IP/ Non-IP
BMS-936558-01 (Nivolumab) Solution for Injection	100 mg (10 mg/mL)	IP
Ipilimumab Solution for Injection	200 mg (5 mg/mL)	IP
Cabozantinib Tablet	20 mg	IP
Sunitinib Malate Capsule	12.5 mg	IP

Abbreviation: IP=investigational product

2. SCHEDULE OF ACTIVITIES

Table 2-1: Screening Procedural Outline (CA2099ER)

Procedure	Screening Visit[a]	Notes
Eligibility Assessments		
Informed Consent	X	Register in Interactive Response Technology (IRT) system to obtain participant number.
Inclusion/Exclusion Criteria	X	Must be confirmed prior to randomization.
Medical History	X	
International Metastatic RCC Database Consortium (IMDC) Prognostic Score	X	See Appendix 6.
Tumor tissue sample (for stratification by PD-L1 tumor expression)	X	Tumor tissue (preferably obtained within 3 months but no more than 12 months prior to enrollment, with an associated pathology report) will be collected. Formalin-fixed paraffin-embedded (FFPE) block or 20 unstained slides: a minimum of 10 slides will be acceptable if tumor tissue is limited. See Section 9.8.2. Central lab will determine PD-L1 tumor expression. Participants must have an evaluable PD-L1 result from the central lab in order to be randomized.
Safety Assessments		
Full Physical Examination, Measurements, Vital Signs, and Performance Status	X	Height, weight, Karnofsky Performance Status (KPS) (Appendix 7), BP, HR, RR, and temperature within 14 days prior to randomization.
Assessment of Signs and Symptoms	X	Within 14 days prior to randomization.
Review of Concomitant Medications	X	Within 14 days prior to randomization.
Serious Adverse Events Assessment	X	Serious Adverse Events from time of consent. See Section 9.2
Electrocardiogram (ECG)	X	Within 28 days prior to randomization. Fridericia corrected QT (QTcF) required. If any time there is an increase in QTcF interval to an absolute value > 500 msec, 2 additional ECGs must be performed each with intervals approximately 3 minutes apart.
Cardiac Ejection Fraction (via Echocardiogram or MUGA)	X	Within 28 days prior to randomization.

Table 2-1: Screening Procedural Outline (CA2099ER)

Procedure	Screening Visit[a]	Notes
Laboratory Tests (includes blood and urine samples)	X	See Section 9.4.1 for additional details on tests required. To be completed locally at each site. Must be performed within 14 days prior to randomization. • CBC w/differential • PT/INR, PTT • Chemistry panel (includes AST, ALT, total bilirubin, ALP, LDH, creatinine, BUN, glucose, albumin, Na, K, Cl, Ca (also corrected), P, Mg, amylase, lipase) • Thyroid panel (includes TSH with free T3 and free T4) • Hepatitis B/C (HBVsAg, HCV antibody or HCV RNA) • HIV if mandated locally (sites in Germany, see Appendix 12) • Urine protein and urine creatinine (for urine protocol/creatinine ratio [UPCR]). If UPCR ≥ 1.0, obtain 24 hour urine protein.
Pregnancy Test	X	WOCBP only. Serum or urine pregnancy test (minimum sensitivity 25 IU/L or equivalent units of HCG) to be done at screening visit and within 24 hours of first dose of study therapy.
Follicle Stimulating Hormone (FSH)	X	For women under the age of 55 years to confirm menopause as needed.
Efficacy Assessments		
Baseline Tumor Assessments	X	CT/MRI of the chest, abdomen, pelvis, brain, and all known sites of disease, performed within 28 days prior to randomization. All scans need to be submitted for blinded independent central review (BICR). See Section 9.1.

Abbreviations: For abbreviations on lab tests refer back to Section 9.4.1.

[a] Some of the assessments referred to in this section may not be captured as data in the eCRF. They are intended to be used as safety monitoring by the treating physician. Additional testing or assessments may be performed as clinically necessary or where required by institutional or local regulations.

Table 2-2: On Treatment Procedural Outline for Arm A Nivolumab Combined with Cabozantinib, Doublet (CA2099ER)

Procedure	Cycle 1, Day 1 Cycle=2 wks	C2 and subsequent visits[a] Each Cycle=2 wks	Notes
Safety Assessments			
Targeted Physical Examination, Vital Signs, Performance Status	X	X	Weight, BP, HR, RR, temperature, and Karnofsky Performance Status (KPS) (Appendix 7). Performance Status to be performed within 72 hours prior to dosing.
Assessment of Signs and Symptoms	X	X	
Adverse Events and Serious Adverse Events Assessment	X	X	Record at each visit. Participants will be followed for drug-related toxicities until these toxicities resolve, return to baseline, or are deemed irreversible. All adverse events will be documented for a minimum of 100 days after last dose (Appendix 3).
Concomitant Medication Review	X	X	Record at each visit.
Electrocardiogram (ECG)	X	X (See notes)	Fridericia corrected QT (OTcF) required. Only for Cycles 1, 4, 7, then every 6 cycles (ie, Cycles 13, 19, 25, etc). If any time there is an increase in QTcF interval to an absolute value > 500 msec, 2 additional ECGs must be performed each with intervals approximately 3 minutes apart.
Laboratory Tests (includes blood and urine samples)	X	X (See notes)	See 9.4.1 for additional details on tests required. Laboratory tests do not need to be repeated at C1D1 if performed within 14 days prior to first dose. After C1D1, within 72 hours prior to dosing: • CBC w/differential at every cycle • Chemistry panel at every cycle (includes AST, ALT, total bilirubin, ALP, LDH, creatinine, BUN, glucose, albumin, Na, K, Cl, Ca, P, Mg) • Amylase and lipase to be done for Cycles, 1, 2, 4, 5, 7, and then every 3 cycles (ie, Cycles 10, 13, 16, etc) • Thyroid panel (includes TSH with reflexive free T3 and free T4) for Cycles 1, 2, 4, 5, 7, and then every 3 cycles (ie, Cycles 10, 13, 16, etc) • Urine protein and urine creatinine (for UPCR, preferred) or urine dipstick for protein every 3 cycles (ie, Cycles 1, 4, 7, 10, etc)

Table 2-2: On Treatment Procedural Outline for Arm A Nivolumab Combined with Cabozantinib, Doublet (CA2099ER)

Procedure	Cycle 1, Day 1 Cycle=2 wks	C2 and subsequent visits[a] Each Cycle=2 wks	Notes
Pregnancy Test	X	X	Within 24 hours prior to the initial administration of study drug, then every 4 weeks ± 7 days. Serum or Urine. WOCBP only.
Efficacy Assessments			
Tumor Assessments	First tumor assessment post-baseline should be performed Week 12 (± 7 days) following randomization. Use same imaging method as was used at baseline. Subsequent tumor assessments should occur at every 6 weeks (± 7 days) until Week 60, then every 12 weeks (± 14 days) until radiographic progression, assessed by the investigator and confirmed by the BICR.		CT/MRI of the chest, abdomen, pelvis, and all known sites of disease. Tumor assessments should be performed at the specified time points regardless of dosing delays. See Section 9.1 for additional details.
Pharmacokinetic (PK)/ Immunogenicity Assessments			
PK blood samples	X	X (See notes)	For details on sampling time points see Table 9.5-1.
Immunogenicity blood samples	X	X (See notes)	For details on sampling time points see Table 9.5-1.
Exploratory Biomarker Assessments			
Whole Blood (DNA) for Genotyping	X		Only prior to dose at Cycle 1. See Sections 9.8.1 and 9.8.6.
Serum Biomarker	X	X (See notes)	Prior to dosing. At Cycles 1, 2, and 4. See Sections 9.8.1 and 9.8.3.

Table 2-2: On Treatment Procedural Outline for Arm A Nivolumab Combined with Cabozantinib, Doublet (CA2099ER)

Procedure	Cycle 1, Day 1 Cycle=2 wks	C2 and subsequent visits[a] Each Cycle=2 wks	Notes
Plasma Biomarkers	X	X (See notes)	Prior to dosing. At Cycles 1, 2, and 4. See Sections 9.8.1 and 9.8.3.
Myeloid Derived Suppressor Cells	X		Prior to dosing. At Cycle 1 only. See Section 9.8.4.
Peripheral blood mononuclear cells (PBMCs)	X	X (See notes)	Prior to dosing: At Cycles 1 and 4. See Sections 9.8.1 and 9.8.5
Tumor Tissue Sample	Every effort should be made to collect fresh tumor tissue sample if available upon disease progression.		If a biopsy has been performed and tissue is available, participants are requested to submit fresh tumor tissue upon disease progression for biomarker research. Tissue submission is optional and biopsy is not required by protocol. See Section 9.8.2.
Participant-Reported Outcomes			
Functional Assessment of Cancer Therapy - Kidney Symptom Index (FKSI-19) Questionnaire	X	X	Completed on Day 1 of each treatment cycle prior to any study-related procedures. See Section 9.1.4.
EuroQoL group's EQ-5D-3L Questionnaire	X	X	Completed on Day 1 of each treatment cycle prior to any study-related procedures. See Section 9.1.4
Health Care Resource Utilization	X	X	See Section 9.9.
Study Treatment			
Randomize	X		Begins with call to Interactive Response Technology (IRT). Participants must have an evaluable PD-L1 result from the central lab in order to be randomized.
Administer Nivolumab and Cabozantinib	X	X	See Section 7. Dispense study treatment as appropriate.

Abbreviations: C=cycle; D=day; wks= weeks. For abbreviations on lab tests refer back to Section 9.4.1.

Notes: Some of the assessments referred to in this section may not be captured as data in the eCRF. They are intended to be used as safety monitoring by the treating physician. Additional testing or assessments may be performed as clinically necessary or where required by institutional or local regulations.

[a] If a dose is delayed, the procedures scheduled for that same timepoint should also be delayed to coincide with when that timepoint's dosing actually occurs.

Table 2-3: On Treatment Procedural Outline for Arm B Nivolumab and Ipilimumab Combined with Cabozantinib, Triplet (CA2099ER)

Procedure	Part 1		Part 2	Notes (Please note differences in cycle durations between Part 1 and Part 2)
	Cycle 1 (Each Cycle=3 wks)	C2-C4[a] (Each Cycle= 3 wks)	C5 and Subsequent visits [a] (Each Cycle=2 wks)	
Safety Assessments				
Targeted Physical Examination, Vital Signs, Performance Status	X	X	X	Weight, BP, HR, RR, temperature, and Karnofsky Performance Status (KPS) (Appendix 7). Performance Status to be performed within 72 hours prior to dosing.
Assessment of Signs and Symptoms	X	X	X	
Adverse Events and Serious Adverse Events Assessment	X	X	X	Record at each visit. Participants will be followed for drug-related toxicities until these toxicities resolve, return to baseline, or are deemed irreversible. All adverse events will be documented for a minimum of 100 days after last dose (Appendix 3).
Concomitant Medication Review	X	X	X	Record at each visit.
Electrocardiogram (ECG)	X	X (See notes)	X (See notes)	Fridericia corrected QT (QTcF) required. Only Cycles 1, 3, 5, then every 6 cycles (ie, Cycles 11, 17, 23, etc). If any time there is an increase in QTcF interval to an absolute value > 500 msec, 2 additional ECGs must be performed each with intervals approximately 3 minutes apart.
Laboratory Tests (includes blood and urine samples)	X	X	X (See notes)	See Section 9.4.1 for additional details on tests required. Laboratory tests do not need to be repeated at C1D1 if performed within 14 days prior to first dose. After C1D1, within 72 hours prior to dosing. • CBC w/differential at every cycle • Chemistry panel at every cycle (includes AST, ALT, total bilirubin, ALP, LDH, creatinine, BUN, glucose, albumin, Na, K, Cl, Ca, P, Mg) • Amylase and lipase at every cycle until Cycle 5 then every 3 cycles (ie, Cycles 8, 11, 14, etc) • Thyroid panel (includes TSH with reflexive free T3 and free T4) at

Table 2-3: On Treatment Procedural Outline for Arm B Nivolumab and Ipilimumab Combined with Cabozantinib, Triplet (CA2099ER)

Procedure	Part 1		Part 2	Notes (Please note differences in cycle durations between Part 1 and Part 2)
	Cycle 1 (Each Cycle=3 wks)	C2-C4[a] (Each Cycle= 3 wks)	C5 and Subsequent visits [a] (Each Cycle=2 wks)	
				every cycle until Cycle 5, then every 3 cycles (ie, Cycles 8, 11, 14, etc) • Urine protein and urine creatinine (for UPCR, preferred) or urine dipstick for protein at Cycles 1, 3, and 5, then every 3 cycles (ie, Cycles 8, 11, 14, etc)
Pregnancy Test	X	X	X	Within 24 hours prior to the initial administration of study drug, then every 4 weeks ± 7 days. Serum or Urine. WOCBP only.
Efficacy Assessments				
Tumor Assessments	First tumor assessment post-baseline should be performed Week 12 (± 7 days) following randomization. Use same imaging method as was used at baseline. Subsequent tumor assessments should occur at every 6 weeks (± 7 days) until Week 60, then every 12 weeks (± 14 days) until radiographic progression, assessed by the investigator and confirmed by the BICR.			CT/MRI of the chest, abdomen, pelvis, and all known sites of disease. Tumor assessments should be performed at the specified time points regardless of dosing delays. See Section 9.1 for additional details.
Pharmacokinetic (PK) /Immunogenicity Assessments				
PK blood samples	X	X (See notes)	X (See notes)	For details on sampling timepoints see Table 9.5-2.
Immunogenicity blood samples	X	X (See notes)	X (See notes)	For details on sampling timepoints see Table 9.5-2.
Exploratory Biomarker Assessments				

Table 2-3: On Treatment Procedural Outline for Arm B Nivolumab and Ipilimumab Combined with Cabozantinib, Triplet (CA2099ER)

Procedure	Part 1		Part 2	Notes (Please note differences in cycle durations between Part 1 and Part 2)
	Cycle 1 (Each Cycle=3 wks)	C2-C4[a] (Each Cycle= 3 wks)	C5 and Subsequent visits [a] (Each Cycle=2 wks)	
Whole Blood (DNA) for Genotyping	X			Only prior to dose at Cycle 1. See Sections 9.8.1 and 9.8.6.
Serum Biomarkers	X	X (See notes)		Prior to dosing. At Cycles 1, 2, and 3. See Sections 9.8.1 and 9.8.3.
Plasma Biomarkers	X	X (See notes)		Prior to dosing. At Cycles 1, 2, and 3. See Sections 9.8.1 and 9.8.3.
Myeloid Derived Suppressor Cells	X			Prior to dosing. At Cycle 1 only. See Section 9.8.4.
Peripheral blood mononuclear cells (PBMCs)	X	X (See notes)		Prior to dosing. At Cycles 1 and 3 only. See Sections 9.8.1 and 9.8.5.
Tumor Tissue Sample	Every effort should be made to collect fresh tumor tissue sample if available upon disease progression.			If a biopsy has been performed and tissue is available, participants are requested to submit fresh tumor tissue upon disease progression for biomarker research. See Section 9.8.2. Tissue submission is optional and biopsy is not required by protocol.
Participant-Reported Outcomes				
Functional Assessment of Cancer Therapy - Kidney Symptom Index (FKSI-19) Questionnaire	X	X	X	Completed on Day 1 of each treatment cycle prior to any study-related procedures. See Section 9.1.4.
EuroQoL group's EQ-5D-3L Questionnaire	X	X	X	Completed on Day 1 of each treatment cycle prior to any study-related procedures. See Section 9.1.4.

Table 2-3: On Treatment Procedural Outline for Arm B Nivolumab and Ipilimumab Combined with Cabozantinib, Triplet (CA2099ER)

Procedure	Part 1		Part 2	Notes (Please note differences in cycle durations between Part 1 and Part 2)
	Cycle 1 (Each Cycle=3 wks)	C2-C4[a] (Each Cycle= 3 wks)	C5 and Subsequent visits [a] (Each Cycle=2 wks)	
Health Care Resource Utilization	X	X	X	See Section 9.9.
Study Treatment				
Randomize	X			Begins with call to Interactive Response Technology (IRT). Participants must have an evaluable PD-L1 result from the central lab in order to be randomized.
Administer Nivolumab, Ipilimumab, and Cabozantinib	X	X		Cycles 1 to 4 are 3 week cycles. See Section 7.
Administer Nivolumab and Cabozantinib			X	Cycle 5 and subsequent cycles are 2 week cycles. See Section 7.
Dispense Study Treatment	X	X	X	See Section 7.

Abbreviations: C=cycle; D=day; wks= weeks. For abbreviations on lab tests refer back to Section 9.4.1.

Notes: Some of the assessments referred to in this section may not be captured as data in the eCRF. They are intended to be used as safety monitoring by the treating physician. Additional testing or assessments may be performed as clinically necessary or where required by institutional or local regulations.

[a] If a dose is delayed, the procedures scheduled for that same timepoint should also be delayed to coincide with when that timepoint's dosing actually occurs.

Table 2-4: On Treatment Procedural Outline for Arm C Sunitinib (CA2099ER)

Procedure	Cycle 1 Each Cycle=6wks		Cycle 2 [a] Each Cycle=6 wks		C3 and Subsequent visits[a] Each Cycle=6 wks		Notes
	Day 1	Day 22 (+/- 3 days)	Day 1	Day 22 (+/- 3 days)	Day 1	Day 22 (+/- 3 days)	
Safety Assessments							
Targeted Physical Examination, Vital Signs, Performance Status	X		X		X		Weight, BP, HR, RR, temperature, and Karnofsky Performance Status (KPS) (Appendix 7). Performance Status to be performed within 72 hours prior to dosing.
Assessment of Signs and Symptoms	X	X	X	X	X	X	
Adverse Events and Serious Adverse Events Assessment	X	X	X	X	X	X	Record at each visit. Participants will be followed for drug-related toxicities until these toxicities resolve, return to baseline, or are deemed irreversible. All adverse events will be documented for a minimum of 100 days after last dose (Appendix 3).
Concomitant Medication Review	X	X	X	X	X	X	Record at each visit
Electrocardiogram (ECG)	X		X		X (See notes)		Fridericia corrected QT (QTcF) required. Only Cycles 1, 2, and 3, then every 2 cycles (ie, Cycles 5, 7, 9, etc). If any time there is an increase in QTcF interval to an absolute value > 500 msec, 2 additional ECGs must be performed each with intervals approximately 3 minutes apart.
Laboratory Tests (includes blood and urine samples)	X	X (See notes)	X	X (See notes)	X	X (See notes)	See Section 9.4.1 for additional details on tests required. Laboratory tests do not need to be repeated at C1D1 if performed within 14 days prior to first dose. After C1D1, within 72 hours prior to specified dosing day: • CBC w/differential on Day 1 and Day 22 (+/- 3 days) of each cycle • Chemistry panel on Day 1 and Day 22 (+/- 3 days) of

Table 2-4: On Treatment Procedural Outline for Arm C Sunitinib (CA2099ER)

Procedure	Cycle 1 Each Cycle=6wks		Cycle 2 [a] Each Cycle=6 wks		C3 and Subsequent visits[a] Each Cycle=6 wks		Notes
	Day 1	Day 22 (+/- 3 days)	Day 1	Day 22 (+/- 3 days)	Day 1	Day 22 (+/- 3 days)	
							every cycle (includes AST, ALT, total bilirubin, ALP, LDH, creatinine, BUN, glucose, albumin, Na, K, Cl, Ca, P, Mg) • Amylase and lipase on Day 1 and Day 22 (+/- 3 days) of Cycle 1 and 2, then on Day 1 of Cycle 3 and all subsequent cycles. • Thyroid panel (includes TSH with reflexive free T3 and free T4) on Day 1 and Day 22 (+/- 3 days) of Cycle 1 and 2, then on Day 1 of Cycle 3 and all subsequent cycles. • Urine protein and urine creatinine (for UPCR, preferred) or urine dipstick for protein on Day 1 of every cycle
Pregnancy Test	X	X	X	X	X	X	Within 24 hours prior to the initial administration of study drug, then every 4 weeks ± 7 days. Serum or Urine. WOCBP only.
Efficacy Assessments							
Tumor Assessments	First tumor assessment post-baseline should be performed Week 12 (± 7 days) following randomization. Use same imaging method as was used at baseline. Subsequent tumor assessments should occur at every 6 weeks (± 7 days) until Week 60, then every 12 weeks (± 14 days) until radiographic progression, assessed by the investigator and confirmed by the BICR.						CT/MRI of the chest, abdomen, pelvis, and all known sites of disease. Tumor assessments should be performed at the specified time points regardless of dosing delays. See Section 9.1 for additional details.
Exploratory Biomarker Assessments							

Table 2-4: On Treatment Procedural Outline for Arm C Sunitinib (CA2099ER)

Procedure	Cycle 1 Each Cycle=6wks		Cycle 2 [a] Each Cycle=6 wks		C3 and Subsequent visits[a] Each Cycle=6 wks		Notes
	Day 1	Day 22 (+/- 3 days)	Day 1	Day 22 (+/- 3 days)	Day 1	Day 22 (+/- 3 days)	
Whole Blood (DNA) for Genotyping	X						Only prior to dose at Cycle 1. See Sections 9.8.1 and 9.8.6
Serum Biomarkers	X		X				Prior to dosing. At Cycles 1 and 2. See Sections 9.8.1 and 9.8.3.
Plasma Biomarkers	X		X				Prior to dosing. At Cycles 1 and 2. See Sections 9.8.1 and 9.8.3.
Myeloid Derived Suppressor Cells	X						Prior to dosing at Cycle 1 only. See Section 9.8.4.
Peripheral blood mononuclear cells (PBMCs)	X		X				Prior to dosing. At Cycles 1 and 2. See Sections 9.8.1 and 9.8.5.
Tumor Tissue Sample	Every effort should be made to collect fresh tumor tissue sample if available upon disease progression.						If a biopsy has been performed and tissue is available, participants are requested to submit fresh tumor tissue upon disease progression for biomarker research. See Section 9.8.2. Tissue submission is optional and biopsy is not required by protocol.
Participant-Reported Outcomes							
Functional Assessment of Cancer Therapy - Kidney Symptom Index (FKSI-19) Questionnaire	X		X		X		Completed on Day 1 of each treatment cycle prior to any study-related procedures. See Section 9.1.4.
EuroQoL group's EQ-5D-3L Questionnaire	X		X		X		Completed on Day 1 of each treatment cycle prior to any study-related procedures. See Section 9.1.4.

Table 2-4: On Treatment Procedural Outline for Arm C Sunitinib (CA2099ER)

Procedure	Cycle 1 Each Cycle=6wks		Cycle 2 [a] Each Cycle=6 wks		C3 and Subsequent visits[a] Each Cycle=6 wks		Notes
	Day 1	Day 22 (+/- 3 days)	Day 1	Day 22 (+/- 3 days)	Day 1	Day 22 (+/- 3 days)	
Health Care Resource Utilization	X		X		X		See Section 9.9.
Study Treatment							
Randomize	X						Begins with call to Interactive Response Technology (IRT). Participants must have an evaluable PD-L1 result from the central lab in order to be randomized.
Administer Sunitinib	X	X	X	X	X	X	Each cycle will be 6 weeks where sunitinib will be administered for 4 weeks, then participants will be off treatment for 2 weeks. See Section 7.
Dispense Study Treatment	X		X		X		See Section 7.

Abbreviations: C=cycle; D=day; wks= weeks. For abbreviations on lab tests refer back to Section 9.4.1.

Notes: Some of the assessments referred to in this section may not be captured as data in the eCRF. They are intended to be used as safety monitoring by the treating physician. Additional testing or assessments may be performed as clinically necessary or where required by institutional or local regulations.

[a] If a dose is delayed, the procedures scheduled for that same timepoint should also be delayed to coincide with when that timepoint's dosing actually occurs.

Table 2-5: Follow-Up Procedural Outline for All Arms (CA2099ER)

Procedure	Safety Follow-up (Follow up Visit 1 (FU1) and Visit 2 (FU2))[a]	Survival Follow-up[b]	Notes
Safety Assessments			
Targeted Physical Examination, Vital Signs, Performance Status	X		Weight, BP, HR, RR, temperature, and Karnofsky Performance Status (KPS) (Appendix 7).
Assessment of Signs and Symptoms	X		
Adverse Events and Serious Adverse Events Assessment	X		Record at each visit. Participants will be followed for drug-related toxicities until these toxicities resolve, return to baseline, or are deemed irreversible. All adverse events will be documented for a minimum of 100 days after last dose (Appendix 3).
Concomitant Medication Review	X		Record at each visit
Electrocardiogram (ECG)	X		Fridericia corrected QT (QTcF) required. If any time there is an increase in QTcF interval to an absolute value > 500 msec, 2 additional ECGs must be performed each with intervals approximately 3 minutes apart.
Laboratory Tests (includes blood and urine samples)	FU1 -yes FU2 - if toxicities are present		See Section 9.4.1 for additional details on tests required. • CBC w/differential, PT/INR, and PTT • Chemistry panel (includes AST, ALT, total bilirubin, ALP, LDH, creatinine, BUN, glucose, albumin, Na, K, Cl, Ca, P, Mg, amylase, lipase) • Thyroid panel (includes TSH with reflexive free T3 and free T4) • Urine protein and urine creatinine (for UPCR, preferred) or urine dipstick for protein
Pregnancy Test	X		Serum or Urine. WOCBP only.
Efficacy Assessments			
Tumor Assessments	First tumor assessment post-baseline should be performed Week 12 (± 7 days) following randomization. Use same imaging method as was used at baseline.		Participants who discontinue study treatment without radiographic progression, confirmed by BICR, will continue tumor assessments according to the protocol specified schedule, even if new anti-tumor therapy has been initiated in the Follow-Up phase, until radiographic

Table 2-5: Follow-Up Procedural Outline for <u>All Arms</u> (CA2099ER)

Procedure	Safety Follow-up (Follow up Visit 1 (FU1) and Visit 2 (FU2))[a]	Survival Follow-up[b]	Notes
	Subsequent tumor assessments should occur at every 6 weeks (± 7 days) until Week 60, then every 12 weeks (± 14 days) until radiographic progression, assessed by the investigator and confirmed by the BICR.		progression has been assessed by the investigator and confirmed by BICR. CT/MRI of the chest, abdomen, pelvis, and all known sites of disease. See Section 9.1.2 for additional details.
Survival Status	X	X	During safety follow up and every 3 months (clinic visit or by telephone) during survival phase. Include documentation of subsequent chemotherapy. See Section 8.1.5.
<u>Pharmacokinetic (PK)/ Immunogenicity Assessments</u>			
PK blood samples	X (See notes)		Only for participants who were in Arm A (Nivolumab combined with cabozantinib, Doublet) or Arm B (Nivolumab and ipilimumab combined with cabozantinib, Triplet). For details on sampling timepoints see Table 9.5-1
Immunogenicity blood samples	X (See notes)		Only for participants who were in Arm A (Nivolumab combined with cabozantinib, Doublet) or Arm B (Nivolumab and ipilimumab combined with cabozantinib, Triplet). For details on sampling timepoints see Table 9.5-1
<u>Participant-Reported Outcomes</u>			
Functional Assessment of Cancer Therapy - Kidney Symptom Index (FKSI-19) Questionnaire	X		See Section 9.1.4
EuroQoL group's EQ-5D-3L Questionnaire	X	X	See Section 9.1.4.
Health Care Resource Utilization	X		See Section 9.9.

Abbreviations: C=cycle; wks= weeks; FU= follow up. For abbreviations on lab tests, see Section 9.4.1.

[a] Participants must be followed for at least 100 days after last dose of study treatment. Follow-up visit #1 (FU1) should occur 30 days from the last dose (+/- 7) days or can be performed on the date of discontinuation if that date is greater than 42 days from last dose. Follow-up visit #2 (FU2) occurs approximately 100 days (+/- 7 days) from last dose of study drug. Both Follow Up visits should be conducted in person.

[b] Survival Follow-up visits to occur every 3 months from Follow-up Visit #2. Survival visits may be conducted in person or by telephone. BMS may request that survival data be collected on all treated participants outside of the 3 month specified window. At the time of this request, each participant will be contacted to determine their survival status unless the participant has withdrawn consent for all contact.

3. INTRODUCTION

Renal cell carcinoma (RCC) is the eighth most common cancer in the world with an increasing incidence.[1] Globally, RCC occurs in more than 330,000 cases with approximately a third of the patients succumbing to their disease. More than 100,000 deaths occur annually, as a result of progression of metastatic disease. Despite the earlier detection of smaller kidney tumors, the rate of RCC-related mortality has increased suggesting that recurrence and advanced disease are responsible for mortality.[2,3] With the rise in RCC incidence, as well as mortality and morbidity associated with advanced RCC, medical need in this population remains a priority.

Over the last decade, an increased understanding of the biology of RCC has led to development of multiple agents that target specific growth pathways. The vascular endothelial growth factor (VEGF) pathway and targeted serine/threonine protein kinase therapies that block the mammalian target of rapamycin (mTOR) have been found to be important targets in RCC disease. Global health authorities (HAs) have approved multiple drugs targeting these pathways, including anti-VEGF agents, such as pazopanib, sorafenib, sunitinib, cabozantinib, and bevacizumab, and mTOR pathway inhibitors, such as temsirolimus and everolimus.[4] Additionally, recent innovation of treating cancer with immunotherapies has also expanded treatment options. Nivolumab, an anti-PD-1 antibody, given as monotherapy or in combination with the anti-CTLA-4 antibody, ipilimumab, has demonstrated clinical activity in multiple tumor types, including RCC.

To better understand treatment and patient outcomes, several academic groups have identified variables associated with survival and created prognostic models in mRCC. These risk models are commonly used for choosing therapies or selecting patients for treatment in clinical trials. The International Metastatic Renal Cell Carcinoma Database Consortium (IMDC) model stratifies patients into 3 prognostic groups, based on 6 adverse prognostic factors, into favorable (0 factors), intermediate (1-2 factors), and poor risk (3-6 factors) groups.[5] The currently available first-line agents for mRCC, which target the VEGF pathway, have shown limited efficacy in the intermediate and poor risk populations, yielding median overall survival of less than 2 years.

Cabozantinib and nivolumab both share category 1 NCCN guideline recommendations (ie, uniform consensus that the treatment is appropriate, based on high-level evidence) for the treatment of previously treated mRCC patients.[6] Therefore, it is an appropriate next step to combine these agents and move them to the first-line setting in an attempt to improve clinical outcomes in patients with advanced RCC. This protocol CA2099ER will test the clinical activity of nivolumab combined with cabozantinib (doublet regimen) or nivolumab and ipilimumab combined with cabozanitinib (triplet regimen). Given the different mechanisms of action of each of these agents, there is potential for distinct improvement in clinical efficacy.[7]

3.1 Study Rationale

Agents that target the VEGF pathway prevent tumor growth by inhibiting angiogenesis. A recent clinical biomarker study has shown that anti-VEGF therapy also affects the RCC tumor immune microenvironment. In this study, RCC tumor tissue from treatment-naive patients and those who

had received prior bevacizumab or sunitinib were assessed for tumor immune cell infiltration. Samples from patients with prior anti-VEGF therapy demonstrated increased infiltration of regulatory T cells (Tregs) and enhanced tumor PD-L1 expression, both of which were negatively associated with patient survival.[8] These effects suggest that the promotion of an immune suppressive tumor microenvironment may contribute to anti-VEGF therapy resistance and point to a rational strategy for combining anti-VEGF therapy with immunotherapies that target the PD-1/PD-L1 pathway. Indeed, the phase 1 study CA209016 demonstrated the synergistic activity of such combinations in mRCC. This study included a cohort of 26 participants with mRCC (19 of whom were previously untreated) who received sunitinib in combination with nivolumab 5 mg/kg Q3W. Using MSKCC risk criteria, 13 participants in this cohort had favorable risk disease, 12 had intermediate risk disease, and 1 had poor risk disease. After a minimum follow-up of 22 months, the objective response rate (ORR) was 42.3%, median progression-free survival (PFS) was 12.3 months, and median overall survival (OS) was 36.8 months in this cohort.[9] Similar efficacy results have been reported in other phase 1 studies which combine anti-VEGF agents with anti-PD-1 agents.[10,11]

As described in Section 3.2, cabozantinib is a novel tyrosine kinase inhibitor that, in addition to VEGFR, targets additional tyrosine kinases that are implicated in the biology of mRCC, such as c-MET and AXL. In a randomized phase 3 trial in patients with advanced RCC that had progressed after anti-VEGFR therapy, cabozantinib was shown to improve PFS and OS compared to everolimus, leading to its regulatory approval. Subsequently, a randomized phase 2 trial of cabozantinib vs sunitinib has demonstrated an improvement in ORR and PFS in intermediate- and poor-risk patients with previously untreated mRCC (see Section 3.2.1.5).

Cabozantinib has also been demonstrated to have effects on immune cells. In a study of 24 subjects with advanced urothelial carcinoma, cabozantinib treatment resulted in a decrease in circulating Tregs and increased PD-1 expression on Tregs. Low Tregs at baseline were also predictive of improved response to cabozantinib and survival.[12]

Given the promising clinical activity of cabozantinib in previously untreated mRCC and its potential immune effects, combining cabozantinib with nivolumab (in a doublet regimen) is a rational strategy to optimize first-line therapy in mRCC. Based on both preclinical[13] and clinical studies that demonstrate the activity of the nivolumab + ipilimumab combination (Section 3.2.1.4), it is also rational to combine nivolumab and ipilimumab with cabozantinib (in a triplet regimen). An ongoing phase 1 study is evaluating both the doublet and triplet regimens in patients with refractory advanced urothelial cancer or other genitourinary tumors, including mRCC, and has defined dosing for both regimens that produces acceptable safety and tolerability (Section 3.2.1.6). This 3-arm randomized phase 3 trial will determine if the combination doublet regimen (nivolumab combined with cabozantinib) and the combination triplet regimen (nivolumab and ipilimumab combined with cabozantinib) produce greater clinical benefit than sunitinib, a standard of care agent for patients with previously untreated mRCC. In addition, this trial will reveal the adverse event profiles, quality of life measures, as well as exploratory biomarkers associated with these different first-line treatment regimens.

3.1.1 Research Hypothesis

Treatment with nivolumab combined with cabozantinib (doublet regimen) or nivolumab and ipilimumab combined cabozantinib (triplet regimen) will demonstrate an improvement in PFS per BICR compared to sunitinib monotherapy in intermediate/poor risk participants with previously untreated mRCC.

3.2 Background

Cancer immunotherapy rests on the premise that tumors can be recognized as foreign rather than as self and can be effectively attacked by an activated immune system. An effective immune response in this setting is thought to rely on immune surveillance of tumor antigens expressed on cancer cells that ultimately results in an adaptive immune response and cancer cell death. Meanwhile, tumor progression may depend upon acquisition of traits that allow cancer cells to evade immunosurveillance and escape effective innate and adaptive immune responses.[14,15,16] Current immunotherapy efforts attempt to break the apparent tolerance of the immune system to tumor cells and antigens by either introducing cancer antigens by therapeutic vaccination or by modulating regulatory checkpoints of the immune system. T-cell stimulation is a complex process involving the integration of numerous positive as well as negative co-stimulatory signals in addition to antigen recognition by the T-cell receptor (TCR).[17] Collectively, these signals govern the balance between T-cell activation and tolerance.

Programed death-1 (PD-1) is a member of the CD28 family of T-cell co-stimulatory receptors that also includes CD28, CTLA 4, ICOS, and BTLA.[18] PD-1 signaling has been shown to inhibit CD-28-mediated upregulation of IL-2, IL-10, IL-13, interferon-γ (IFN-γ) and Bcl-xL. PD-1 expression also been noted to inhibit T cell activation, and expansion of previously activated cells. Evidence for a negative regulatory role of PD-1 comes from studies of PD-1 deficient mice, which develop a variety of autoimmune phenotypes.[19] These results suggest that PD-1 blockade has the potential to activate anti-self T-cell responses, but these responses are variable and dependent upon various host genetic factors. Thus, PD-1 deficiency or inhibition is not accompanied by a universal loss of tolerance to self-antigens.

In vitro, nivolumab (BMS-936558) binds to PD-1 with high affinity (EC_{50} 0.39-2.62 nM), and inhibits the binding of PD-1 to its ligands PD-L1 and PD-L2 (IC_{50} ± 1 nM). Nivolumab binds specifically to PD-1 and not to related members of the CD28 family such as CD28, ICOS, CTLA-4 and BTLA. Blockade of the PD-1 pathway by nivolumab results in a reproducible enhancement of both proliferation and IFN-γ release in the mixed lymphocyte reaction (MLR). Using a CMV re stimulation assay with human PBMC, the effect of nivolumab on antigen specific recall response indicates that nivolumab augmented IFN-γ secretion from CMV specific memory T cells in a dose-dependent manner versus isotype-matched control. In vivo blockade of PD-1 by a murine analog of nivolumab enhances the anti-tumor immune response and result in tumor rejection in several immunocompetent mouse tumor models (MC38, SA1/N, and PAN02).[20]

CTLA-4, an activation-induced T-cell surface molecule, is a member of the CD28:B7 immunoglobulin superfamily that competes with CD28 for B7. CTLA-4 mediated signals are inhibitory and turn off T cell-dependent immune responses.[21] Ipilimumab is a fully human monoclonal IgG1κ that binds to the CTLA-4 antigen expressed on a subset of T cells from human and nonhuman primates. The proposed mechanism of action for ipilimumab is interference of the interaction of CTLA-4 with B7 molecules on antigen presenting cells, with subsequent blockade of the inhibitory modulation of T-cell activation promoted by the CTLA 4/B7 interaction.

Cabozantinib inhibits multiple receptor tyrosine kinases (RTKs) implicated in tumor growth, metastasis, and angiogenesis.[22] The primary targets of cabozantinib implicated in mRCC are MET (c-MET), AXL and vascular endothelial growth factor receptor 2 (VEGFR2); additional targets identified in-vitro include RET, KIT, ROS1, TYRO3, MER, KIT, TRKB, FLT-3, and TIE-2. Both c-Met, AXL, and VEGFR2 are important mediators of tumor growth and tumor angiogenesis, and in vivo pharmacodynamic activity of cabozantinib against c-Met, and VEGFR2 has been demonstrated in both preclinical and clinical studies.[23,24,25,26]

In addition, preclinical and clinical observations have suggested that cabozantinib promotes an immunopermissive environment which might present an opportunity for synergistic effects from combination treatment with PD-1 checkpoint inhibitors. Specifically, treatment of tumor cells with cabozantinib in vitro led to increased tumor-cell expression of major histocompatibility complex (MHC) class 1 antigen and greater sensitivity of tumor cells to T-cell-mediated killing.[27] In a mouse tumor model, cabozantinib treatment led to increased peripheral CD8+ T-cell counts, decreased regulatory T-cells (Tregs) and myeloid-derived suppressor cells (MDSCs), and decreased Treg suppressor activity. Further, synergistic effects including increased CD8+ T-cell infiltration and decreased infiltration by MDSCs and tumor-assisted macrophages (TAMs) were observed when a poxviral-based cancer vaccine was administered in addition to cabozantinib in the mouse tumor model. Reductions in immunosuppressive Treg lymphocytes following treatment with cabozantinib have also been observed in subjects with advanced refractory urothelial cancer.[12] In a phase 2 study in metastatic triple-negative breast cancer, cabozantinib-treated subjects experienced a persistent increase in the fraction of circulating CD3+ T lymphocytes and a persistent decrease in the CD14+ monocytes possibly reflecting activation of systemic antitumor immunity.[28]

A detailed description of the chemistry, pharmacology, efficacy, and safety of nivolumab and ipilimumab in combination and cabozantinib is provided in the Nivolumab Investigator Brochure (IB) and Cabozantinib IB, respectively.[22,29,30]

3.2.1 Indication Background

3.2.1.1 Sunitinib in Renal Cell Carcinoma

Sunitinib is a VEGF receptor TKI that is approved and recommended for the treatment of mRCC across prognostic groups.[6,31] In a randomized phase 3 trial of sunitinib vs IFNα in treatment-naive subjects (including 36% with favorable risk, 57% with intermediate risk, and 7%

with poor risk per MSKCC criteria), PFS (by independent radiology review) was significantly improved in the sunitinib group compared to the IFNα group (median PFS 11 vs 5 months, HR = 0.42; p < 0.001). ORR was also greater in the sunitinib group (31% vs 6%). Median OS was 26.4 months in the sunitinib group vs 21.8 months in the IFNα group (HR = 0.82; p = 0.051).[32] More recently, sunitinib was compared to pazopanib in treatment-naive subjects (including 27% favorable risk, 73% intermediate risk, and no poor risk) in the phase 3 COMPARZ study.[33] In this non-inferiority study, sunitinib and pazopanib demonstrated similar median PFS (8.4 months for pazopanib vs 9.5 months for sunitinib, HR = 1.05) and median OS (28.4 months for pazopanib vs 29.3 months for sunitinib, HR = 0.91, p = 0.28). The ORR of pazopanib and sunitinib was 31% and 24%, respectively. The most common (≥ 20% frequency) adverse reactions include fatigue, asthenia, fever, diarrhea, nausea, mucositis/stomatitis, vomiting, dyspepsia, abdominal pain, constipation, hypertension, peripheral edema, rash, hand-foot syndrome, skin discoloration, dry skin, hair color changes, altered taste, headache, back pain, arthralgia, extremity pain, cough, dyspnea, anorexia, and bleeding.[34] Other important adverse reactions include hepatotoxicity, QT prolongation (including Torsades de Pointes), osteonecrosis of the jaw, tumor lysis syndrome, and thyroid dysfunction

3.2.1.2 Nivolumab Monotherapy in Renal Cell Carcinoma

Harnessing the immune system would be an attractive opportunity, together with efforts to find cell surface markers that can be used to trace and target dormant renal-cell carcinoma cells. Nivolumab monotherapy has been studied in participants with advanced RCC in several BMS-sponsored studies (phases 1 through 3): MDX1106-03, CA209009, CA209010, and CA209025. MDX1106-03 was a phase 1 refractory solid tumor trial, which included 34 participants with previously-treated advanced RCC who received nivolumab at 1 mg/kg (n = 18) or 10 mg/kg (n = 16) given every 2 weeks.[25,35,36] Median progression-free survival (PFS) was 7.3 months. In both the 1 mg/kg and 10 mg/kg cohorts, approximately 30% of participants experienced an objective response with median duration of response of 12.9 months. Responses were generally rapid with a median time to response of 16 weeks. Notably, responses could occur after treatment cessation and persist off treatment. Median overall survival (mOS) was 22.4 months. These results were promising given that many of the participants were heavily pre-treated with 71% having had 2 or more lines of therapy. Treatment-related adverse events (AEs) of any grade were observed in 85% of RCC participants, the most common being fatigue (41%), rash (27%), diarrhea (18%), and pruritus (18%). Grade 3-4 treatment-related AEs were observed in 18% of RCC participants. The spectrum, frequency, and severity of treatment-related AEs were similar in the RCC population compared to the overall study population and were similar across dose levels.

In order to identify a potential dose-response relationship in RCC, a randomized phase 2 study (CA209010) was conducted in 168 participants with advanced RCC previously treated with an antiangiogenic therapy who received nivolumab at 0.3 mg (n = 60), 2 mg/kg (n = 24) or 10 mg/kg (n=54) given every 3 weeks.[37] No dose response relationship was found as measured by PFS, with median PFS of 2.7, 4.0, and 4.2 months for the 0.3, 2, and 10 mg/kg groups, respectively (P = 0.9). ORR was 20%, 22%, and 20% in the 0.3, 2, and 10 mg/kg groups,

respectively. Median time to achievement of an objective response was 2.8-3.0 months. The median duration of response was 22.3 months (4.8, NR) in the 10 mg/kg arm and not yet reached in the 2 lower dose cohorts. Median OS was at 18.2 to 25.5 months, with a minimum of follow-up of 24 months. Fatigue was the most frequent toxicity (22-35%). No new toxicities were identified with 11% experiencing Grade 3-4 treatment-related AEs, none of which were due to pneumonitis. Treatment-related AEs led to discontinuation of study drug in 7% of participants.

A parallel biomarker-focused trial, CA209009, using the same 3 nivolumab dose levels (03, 2, and 10 mg/kg every 3 weeks) was executed to explore predictors of response and identify mechanisms of resistance.[38] This study included 67 participants with previously-treated, advanced RCC who were randomized to one of the 3 nivolumab dose groups and 24 participants with previously-untreated RCC who received nivolumab at 10 mg/kg every 3 weeks. The results mirrored the efficacy and toxicity profile of CA209010 with an overall ORR of 18% in previously-treated participants, and 13% in previously untreated participants and disease stabilization in another 32% of previously treated and untreated participants. At 24 weeks, 36% of participants were free from progression. Of 56 participants with evaluable pretreatment tumor samples, 18 (32%) had ≥ 5% PD-L1 tumor expression. ORR was 22% among those with ≥ 5% PD-L1 tumor expression versus 8% among those with < 5% PD-L1 tumor expression.

Based on the clinical activity of nivolumab observed in these phase 1 and 2 studies, a large phase 3 trial (CA209025) was conducted in 821 participants with advanced RCC previously treated with 1 or 2 antiangiogenic therapies who were randomized to receive nivolumab 3 mg/kg every 2 weeks or everolimus 10 mg daily. A planned interim analysis, after a minimum of follow-up of 14 months, demonstrated a statistically significant and clinically meaningful improvement in OS of nivolumab monotherapy vs everolimus (median OS, 25.0 months vs 19.6 months, respectively; HR 0.73 [98.5% CI, 0.57 to 0.93, P = 0.002). ORR was 25% for nivolumab vs 5% for everolimus. Additional efficacy results are presented in Table 3.2.1.2-1. Among 756 participants with quantifiable PD-L1 tumor expression in pretreatment samples, 24% had ≥1% PD-L1 expression. Among participants with ≥1% PD-L1 expression, median OS was 21.8 months in the nivolumab group and 18.8 months in the everolimus group (HR, 0.79; 95% CI, 0.53 to 1.17). Among participants with <1% PD-L1 expression, the median OS was 27.4 months in the nivolumab group and 21.2 months in the everolimus group (HR, 0.77; 95% CI 0.60 to 0.97). No new safety concerns were identified, and nivolumab monotherapy showed a favorable safety profile as compared to everolimus, evidenced by the lower rates of drug-related AEs (all grades, 79% vs 88%; Grade 3-4, 19%-37%, respectively) and drug-related AEs leading to discontinuation (all grades, 8% vs 13%, respectively) in the nivolumab group. These results were the basis for regulatory approval of nivolumab monotherapy in advanced RCC.

Table 3.2.1.2-1: Summary of Efficacy Results - All Randomized Subjects - CA209025

Efficacy Parameters	Nivolumab (N = 410)	Everolimus (N = 411)
Primary Endpoint		
Overall Survival		
Events , n (%)	183/410 (44.6)	215/411 (52.3)
Stratified log-rank test P value[a,b]	0.0018	
HR (98.52% CI)[c]	0.73 (0.57, 0.93)	
Median (95% CI), months[d]	25.00 (21.75, NR)	19.55 (17.64, 23.06)
Rate at 6 months (95% CI), %[d]	89.2 (85.7, 91.8)	81.2 (77.0, 84.7)
Rate at 12 months (95% CI), %[d]	76.0 (71.5, 79.9)	66.7 (61.8, 71.0)
Secondary Endpoints		
Objective Response Rate per Investigator (CR + PR)[e]		
N (%)	103 (25.1)	22 (5.4)
95% CI[f]	(21.0, 29.6)	(3.4, 8.0)
Odds ratio estimate (95%CI)[g,h]	5.98 (3.68, 9.72)	
P Value[i]	< 0.0001	
Duration of response[e]		
Ongoing responders, n/N (%)	49/103 (47.6)	10/22 (45.5)
Median (95% CI), months[d]	11.99 (7.85, 23.03)	11.99 (6.44, NR)
Min, Max[j]	0.0, 27.6+	0.0+, 22.2+
Progression-free survival		
Events, n (%)	318 (77.6)	322 (78.3)
Stratified log-rank test p value[a]	0.1135	
HR (95% CI)[c]	0.88 (0.75, 1.03)	
Median 95% CI)	4.60 (3.71, 5.39)	4.44 (3.71, 5.52)

[a] Log-rank test stratified by the MSKCC risk group (poor vs intermediate vs favorable), the number of prior antiangiogenic therapies in the advanced/metastatic setting (1 vs 2), and the region (W. Europe, US/Canada vs Rest of the World) as entered into the IVRS.

[b] Based on the 398 observed deaths and O'Brien-Fleming alpha spending function, the boundary for statistical significance requires the P value to be less than 0.0148.

[c] Stratified Cox proportional hazard model. Hazard ratio is nivolumab over everolimus.

[d] Based on Kaplan-Meier Estimates.

[e] The confirmed ORR was 88/410 (21.5%) in the nivolumab group and 16/411 (3.9%) in the everolimus group (stratified CMH test P value < 0.0001), with a median DOR of 23.03 months in the nivolumab group and 13.73 months in the everolimus group.

[f] CR+PR, confidence interval based on the Clopper and Pearson method.

[g] Cochran-Mantel-Haenzel test stratified by the MSKCC risk group (poor vs intermediate vs favorable), the number of prior anti-angiogenic therapies in the advanced/metastatic setting (X vs 2) and the region (Western Europe vs US/Canada vs Rest of the World) as entered into the IVRS.

[h] Ratio of nivolumab over everolimus

[i] Two-sided p value from CMH test for the comparison of the odds ratio of nivolumab over everolimus.

[j] Symbol + indicated a censored value.

3.2.1.3 Ipilimumab in Renal Cell Carcinoma

Ipilimumab monotherapy for the treatment of mRCC was studied in the phase 2 clinical trial MDX010-11.[39] Two sequential cohorts were studied, each with a loading dose of 3 mg/kg followed by 3 doses of either 1 mg/kg (group 3-1; n = 21) or 3 mg/kg (group 3-3; n = 40). Participants with stable disease or partial or complete response were allowed additional treatment. In Group 3-1 (n = 21), 1 participant (5%) had a PR.[26] In Group 3-3 (n = 40), 5 participants (12.5%) had a PR. Among 14 treatment-naive participants in Group 3-3, 3 (21%) had a PR.

In the ipilimumab monotherapy phase 2 clinical trial MDX010-11, the major toxicities were colitis (all Grade 3 & 4; 14% in Group 3-1, 33% in Group3-3) and hypophysitis (1 Grade 3/4, 1 Grade 1/2 in Group 3-3; none in Group 3-1). Most reported AEs were Grade 1/2 (57% in Group 3-1, 35% in Group 3-3) or Grade 3 (38% in Group 3-1, 48 % in Group 3-3).[40] There were 6 participants (15%) with Grade 4 AEs in Group 3-3. The most common treatment-related AEs in Group 3-1 (total 81%) and Group 3-3 (total 93%) were diarrhea (38% and 40%, respectively) and fatigue (33% and 38%, respectively). Most AEs were manageable with appropriate treatment, including high dose corticosteroids and hormone replacement.

3.2.1.4 Nivolumab Plus Ipilimumab in Renal Cell Carcinoma

Promising safety and efficacy results were also observed with the combination of nivolumab and ipilimumab in the advanced RCC population in study CA209016,[9] a phase 1 dose-escalation study of nivolumab in combination with VEGFR-TKIs or ipilimumab in participants with metastatic RCC. Treatment-experienced and -naive participants with metastatic RCC were randomized to receive nivolumab 3 mg/kg + ipilimumab 1 mg/kg (arm N3 + I1) or nivolumab 1 mg/kg + ipilimumab 3 mg/kg (arm N1 + I3) IV Q3W for 4 doses then nivolumab 3 mg/kg IV Q2W until progression/toxicity. In Arm N1+I3, 25 out of 47 participants (53%) were treatment-naive, and 61.7% were intermediate risk according to MSKCC criteria, and 4.3% were categorized as poor risk. In Arm N1 + I3, 21 of 47 participants (45%) were treatment-naive, and 66.0% were in the intermediate risk category, and 6.4% were categorized as poor risk. The primary objective was to assess safety/tolerability; secondary objective was to assess antitumor activity.

After a minimum of 22 months, the level of clinical activity, as measured by confirmed ORR, for the combination of nivolumab and ipilimumab in CA209016 was substantially greater than that

observed in studies of either nivolumab monotherapy (Section 3.2.1.2) or ipilimumab monotherapy (Section 3.2.1.3) in metastatic RCC, including in the treatment-naive subpopulation. The dosing regimen including nivolumab 3 mg/kg combined with ipilimumab 1 mg/kg (N3 + I1) was chosen for further clinical evaluation because it exhibited similar clinical activity to nivolumab 1 mg/kg combined with ipilimumab 3 mg/kg (N1 + I3) but had a more favorable safety profile.

Table 3.2.1.4-1: Antitumor Activity in All Participants (CA209016)[9]

	N3 + I1 (n = 22) Previously Treated	N3 + I1 (n = 25) Treatment-Naive	N1 + I3 (n = 26) Previously Treated	N1 + I 3 (n = 21) Treatment-naive
Confirmed ORR, n (%) (95% CI)	10 (45.5)	9 (36.0) 18.0, 57.5)	10 (38.5)	9 (42.9) (21.8, 66.0)
Median duration of response, weeks (range)	60.1 (9.29, NA)	88.7 (30.00, 105.00)	74.4 (12.29, 108.29)	5 (55.6)
Ongoing responses, % (n/N)	4 (40.0)	4 (44.4)	2 (20.0)	NR (23.57, NA)
Best objective response, n (%)				
Complete response	3 (13.6)	2 (8.0)	0	0
Partial response	7 (31.8)	7 (28.0)	10 (38.5)	9 (42.9)
Stable disease	6 (27.3)	13 (52.0)	11 (42.3)	6 (28.8)
Progressive disease	6 (27.3)	6 (27.3)	3 (11.5)	5 (23.8)
Unable to determine	0	0	2 (7.7)	1 (4.8)
PFS Median months (CI)	6.6 (1.41, 16.39)	8.3 (3.55, 19.29)	10.1 (5.42, 20.76)	8.5 (2.00, NA)
6-month PFS Median months (CI)	54.5 (32.1, 72.4)	56.5 (34.3, 73.8)	65.4 (44.0, 80.3)	61.9 (38.1, 78.8)
Median OS	NR (10.02, NA)	NR (26.68, NA)	30.9 (25.99, NA)	NR (17.45, NA)

Abbreviations: NR = Not Reached

Among the 91 participants treated with nivolumab + ipilimumab combination in CA209016 who provided evaluable baseline tumor samples, 37.4% had ≥ 1% PD-L1 tumor expression, and 16.5% had ≥5% PD-L1 tumor expression. ORR was 47.1% among participants with ≥1% PD-L1 expression and 36.8% among participants with < 1% PD-L1 expression. Among participants with ≥5% PD-L1 expression, ORR was 40.0%.[9]

Among all treated patients in the Arms N3 + I1 and N1 + N3, AEs were seen in 43/47 (91.5%) participants in the N3 + I1 arm and 45/47 (95.7%) participants in the N1 + I3 arm. In the N1 + I3 arm, the most frequently reported drug-related AEs were fatigue (51.1%); rash, and pruritus (each 31.9%); nausea (27.7%); arthralgia (25.5%). In the N1 + I3 arm, the most frequently reported drug related AEs were fatigue (68.1%); diarrhea, and nausea (each 44.7%); pruritus (36.2%); lipase increased (34%); AST increased (31.9%); ALT increased, and decreased appetite

(29.8%); hypothyroidism (27.7%); and rash (25.5%). In the N3 + I1 arm, the most frequently reported, Grade 3-4 drug-related AE was lipase increased (14.9%). In the N1 + I3 arm, the most frequently reported Grade 3-4 drug related AEs were lipase increased (27.7%); ALT increased (21.3%); diarrhea, and colitis (14.9%); AST increased (12.8%).[9]

Treatment-related AEs (including Grade 3-4), treatment-related AEs leading to discontinuation, and treatment-related SAEs all occurred more commonly in participants in the N1 + I3 arm than in the N3 + I1 arm (Table 3.2.1.4-2).[29]

Table 3.2.1.4-2: Summary of Safety Results - All Treated Subjects

	Arm I-1 IPI1 + NIV3 N = 47		**Arm I-3 IPI3 + NIV1 N = 47**	
Death, n (%)	16 (34.0)		18 (38.3)	
Within 30 Days of Last Dose	0		1 (2.1)	
Within 100 Days of Last Dose	3 (6.4)		4 (8.5)	
Due to Study Drug Toxicity	0		0	
	Any Grade	**Grade 3-4**	**Any Grade**	**Grade 3-4**
All-causality SAEs, n (%)	29 (61.7)	20 (42.6)	30 (63.8)	24 (51.0)
Drug-related SAEs, n (%)	11 (23.4)	9 (19.1)	16 (34.0)	16 (34.0)
All-causality AEs Leading to Discontinuation, n (%)	5 (10.6)	3 (6.4)	15 (31.9)	11 (23.4)
Drug-related AEs Leading to Discontinuation, n (%)	5 (10.6)	3 (6.4)	13 (27.7)	9 (19.1)
All-causality AEs, n (%)	47 (100.0)	33 (70.2)	47 (100.0)	34 (72.3)
Drug-related AEs, n (%)	43 (91.5)	18 (38.3)	45 (95.7)	29 (61.7)
All-causality Select AEs, within 30 Days of Last Dose, by Category, n (%)	**Any Grade**	**Grade 3-4**	**Any Grade**	**Grade 3-4**
Endocrine	14 (29.8)	3 (6.4)	19 (40.4)	0
Gastrointestinal	16 (34.0)	3 (6.4)	25 (53.2)	12 (25.5)
Hepatic	11 (23.4)	3 (6.4)	15 (31.9)	8 (17.0)
Pulmonary	3 (6.4)	0	5 (10.6)	0
Renal	11 (23.4)	2 (4.3)	10 (21.3)	2 (4.3)
Skin	29 (61.7)	1 (2.1)	33 (70.2)	1 (2.1)
Hypersensitivity/Infusion Reactions	5 (10.6)	0	3 (6.4)	0

Table 3.2.1.4-2: Summary of Safety Results - All Treated Subjects

	Arm I-1 IPI1 + NIV3 N = 47		Arm I-3 IP13 + NIV1 N = 47	
Drug-related Select AEs, within 30 Days of Last Dose, by Category, n (%)	**Any Grade**	**Grade 3-4**	**Any Grade**	**Grade 3-4**
Endocrine	13 (27.7)	2 (4.3)	19 (40.4)	0
Gastrointestinal	12 (25.5)	2 (4.3)	21 (44.7)	11 (23.4)
Hepatic	9 (19.1)	3 (6.4)	13 (27.7)	8 (17.0)
Pulmonary	3 (6.4)	0	5 (10.6)	0
Renal	9 (19.1)	2 (4.3)	6 (12.8)	1 (2.1)
Skin	23 (48.9)	0	28 (59.6)	1 (2.1)
Hypersensitivity/Infusion Reactions	5 (10.6)	0	3 (6.4)	0
All-causality Immune-mediated AEs, by Category				
Immune-mediated AEs Treated with Immune-modulating medication	**Any Grade**	**Grade 3-4**	**Any Grade**	**Grade 3-4**
Diarrhea/Colitis	3 (6.4)	2 (4.3)	12 (25.5)	10 (21.3)
Hepatitis	5 (10.6)	2 (4.3)	11 (23.4)	8 (17.0)
Pneumonitis	1 (2.1)	0	5 (10.6)	0
Nephritis and Renal Dysfunction	2 (4.3)	1 (2.1)	1 (2.1)	0
Rash	8 (17.0)	1 (2.1)	9 (19.1)	1 (2.1)
Hypersensitivity	0	0	0	0
Immune-Mediated Endocrine AEs Treated with or without Immune-Modulating Medications	**Any Grade**	**Grade 3-4**	**Any Grade**	**Grade 3-4**
Adrenal Insufficiency	3 (6.4)	1 (2.1)	6 (12.8)	0
Hypophysitis	1 (2.1)	1 (2.1)	2 (4.3)	0
Hypothyroidism/Thyroiditis	10 (21.3)	0	14 (29.8)	0
Hyperthyroidism	4 (8.5)	1 (2.1)	8 (17.0)	0
Diabetes Mellitus	0	0	0	0

MedDRA version 18.1; CTC version 4.0. All events are within 100 days of the last dose of study drug, unless otherwise indicated. Sources: Table S.6.2A (deaths), Table S.6.3A (all-causality SAEs), Table S.6.3B (drug-related SAEs), Table S.6.4B (all-causality AEs leading to discontinuation), Table S.6.4D (drug-related AEs leading to discontinuation), Table S.6.2 (all-causality AEs), Table S.6.3.1 (drug-related AEs), Table S.6.101 (all-causality select AEs), Table S.6.105 (all-causality endocrine select AEs), Table S.6.103 (drug-related select AEs), Table S.6.107 (drug-related endocrine select AEs), Table S.6.202 (all-causality IMAEs with exception of endocrine), and Table S.6.204 (all-causality endocrine IMAEs)

A large phase 3 trial, study CA209214, is currently ongoing to determine if nivolumab combined with ipilimumab (N3+ I1 regimen) improves PFS, OS, and ORR vs sunitinib, in participants with previously untreated, advanced or metastatic RCC.

A detailed description of the chemistry, pharmacology, efficacy, and safety of nivolumab and ipilimumab in combination is provided in the Nivolumab Investigator Brochure and Product Label.[29] Note that the nivolumab and ipilimumab combination is currently approved for unresectable or metastatic melanoma, using the N1 + I3 regimen.

3.2.1.5 Cabozantinib Monotherapy in Renal Cell Carcinoma

Cabozantinib is a small molecule inhibitor of the tyrosine kinases c-Met, AXL, and VEGFR2, and has been shown to reduce tumor growth, metastasis, and angiogenesis. Cabozantinib has been evaluated in both first-line and second-line settings in advanced and metastatic RCC.

In participants with advanced renal cell carcinoma (RCC) who progressed after previous VEGFR tyrosine-kinase inhibitor (VEGFR-TKI) treatment, the randomized phase 3 METEOR trial compared the efficacy and safety of cabozantinib monotherapy at a daily dose of 60 mg versus the mTOR inhibitor everolimus at a daily dose of 10 mg.[41,42] In the trial, 658 participants were randomized to receive cabozantinib (n = 330) or everolimus (n = 328). After a minimum follow-up of 11 months in the first 375 randomized subjects, the primary endpoint of PFS (by independent radiology review) was 7.4 months in the cabozantinib arm vs 3.8 months in the everolimus arm (HR 0.58 [95% CI 0.45 to 0.75, p<0.001). Using Memorial Sloan Kettering Cancer Center (MSKCC) criteria, subgroup analyses of risk groups demonstrated PFS benefit in participants with favorable risk (HR 0.54 [95% CI 0.37, 0.79]), intermediate risk (HR 0.56 [95% CI 0.37, 0.84]), and poor risk (HR 0.84 [95% CI 0.46, 1.53]). ORR was 17% in the cabozantinib arm vs 5% in the everolimus arm (p<0.001). An interim OS analysis at the time of this final PFS analysis demonstrated a trend toward longer OS (HR 0.67, p=0.005, with p≤0.0019 required for statistical significance). A subsequent OS analysis, performed after a median follow-up of approximately 19 months in all 658 randomized subjects, demonstrated a significant improvement in OS, with median OS of 21.4 months in the cabozantinib arm and 16.5 months in the everolimus arm (HR 0.66 [95% CI 0.53–0.83]; p = 0.00026). Subgroup analyses of OS according to MSKCC risk group were consistent with the results for the overall population. The most common grade 3 or 4 adverse events included hypertension (49 [15%] in the cabozantinib group vs 12 [4%] in the everolimus group), diarrhea (43 [13%] vs 7 [2%]), fatigue (36 [11%] vs 24 [7%]), palmar-plantar erythrodysesthesia syndrome (27 [8%] vs 3 [1%]), anemia (19 [6%] vs 53 [17%]), hyperglycemia (3 [1%] vs 16 [5%]), and hypomagnesemia (16 [5%] vs none). Serious adverse events grade ≥ 3 occurred in 130 (39%) participants in the cabozantinib group and in 129 (40%) in the everolimus group. One treatment-related death occurred in the cabozantinib group (death; not otherwise specified) and two occurred in the everolimus group (1 aspergillus infection and 1 pneumonia aspiration).[41,42]

In treatment naive participants, cabozantinib was also evaluated in a randomized phase 2 multicenter trial against sunitinib as first-line therapy in participants with advanced or metastatic RCC.[43] Participants were required to have either intermediate or poor risk disease according to IMDC criteria and were randomized in 1:1 ratio to cabozantinib (n = 79) or sunitinib (n = 78). Investigator-assessed PFS was the primary endpoint. Compared with sunitinib, cabozantinib treatment significantly increased median PFS (8.2 v 5.6 months) and was associated with a 34%

reduction in rate of progression or death (adjusted hazard ratio, 0.66; 95% CI, 0.46 to 0.95; one-sided P = .012). ORR was 46% (95% CI, 34 to 57) for cabozantinib vs 18% (95% CI, 10 to 28) for sunitinib. Median OS was 30.3 months for cabozantinib vs 21.8 months for sunitinib (adjusted HR 0.80, 95% CI 0.50, 1.26). All-causality grade 3 or 4 adverse events were 67% for cabozantinib and 68% for sunitinib and included diarrhea (cabozantinib, 10% v sunitinib, 11%), fatigue (6% v 15%), hypertension (28% v 22%), palmar-plantar erythrodysesthesia (8% v 4%), and hematologic adverse events (3% v 22%). Treatment-related Grade 5 events occurred in 3 participants in the cabozantinib arm (acute kidney injury, sepsis, and jejunal perforation) and 3 participants in the sunitinib arm (sepsis, respiratory failure, and vascular disorders).

3.2.1.6 Nivolumab Combined with Cabozantinib or Nivolumab and Ipilimumab Combined with Cabozantinib

The ongoing phase 1 trial (NCT02496208) evaluating the safety and efficacy of nivolumab combined with cabozantinib (doublet regimen) and nivolumab and ipilimumab combined with cabozantinib (triplet regimen) in participants with refractory metastatic urothelial carcinoma (mUC) and other genitourinary tumors has reported interim results.[44] The primary objective was to determine the dose limiting toxicity and recommended phase 2 dose of the doublet and triplet regimens.

At the time of an update in February 2017, 48 participants (enrolled from 22-Jul-2015 to 31-Dec-2016) had been treated, including 19 with urothelial carcinoma; 9 with castration-resistance prostate cancer; 4 with urachal adenocarcinoma; 4 with germ cell tumor; 4 with penile cancer; 2 with squamous cell carcinoma of the bladder/urethra; 2 with clear cell RCC; 2 with sarcomatoid RCC; 1 with trophoblastic tumor; and 1 with Sertoli cell tumor.[44] Part 1 of the study was the dose escalation for the doublet regimen and enrolled first, followed by Part 2, which was the dose escalation for the triplet regimen (which dosed nivolumab and ipilimumab Q3W for the first 4 doses, followed by nivolumab alone Q2W). Thirty participants were treated with the doublet, and 18 were treated with triplet. Both the doublet and triplet regimens were found to be safe and tolerable. The recommended phase 2 dose for the doublet was cabozantinib 40 mg/day + nivolumab 3 mg/kg. The recommended phase 2 dose for the triplet was cabozantinib 40 mg/day + nivolumab 3 mg/kg + ipilimumab 1 mg/kg.

In the doublet arm, the most common treatment-related AEs of any grade were ALT increased (67%); fatigue (63%); diarrhea (60%); hypothyroidism (57%); ALT increased (50%); and anorexia (47%). The most common treatment-related Grade 3-4 AEs were neutropenia (17%); hypophosphatemia and lipase increase (13% each); hypertension (10%); fatigue, diarrhea, thrombocytopenia, and dehydration (7% each). One participant developed immune-mediated Grade 3 aseptic meningitis.

In the triplet arm, the most common treatment-related AEs of any grade were fatigue (72%); diarrhea and anorexia (61% each); ALT increased, hypophosphatemia, dysgeusia, and lipase increased (44% each). The most common treatment-related Grade 3-4 AEs were hypertension and hypophosphatemia (17% each); and fatigue, hyponatremia, nausea, and lipase increase (13% each). One participant developed immune-mediated Grade 3 colitis.

Among 43 participants who were evaluable for response, the ORR was 30%. The ORR was 38% among 26 evaluable participants treated with the doublet and 18% among 17 participants treated with the triplet. One of 2 participants with sarcomatoid RCC achieved a response. Neither of the 2 participants with clear cell RCC were evaluable for response at the time of the analysis.

The doublet and triplet expansion cohorts are continuing to enroll participants with advanced urothelial carcinoma and RCC. Updated safety and efficacy results after additional enrollment and longer follow-up are awaited.

3.3 Benefit/Risk Assessment

The currently available first-line agents for treatment of mRCC are associated with median OS of 43.2 months in patients with favorable risk disease according to IMDC criteria.[45] However, median OS is 22.5 months in those with intermediate risk disease and only 7.8 months in those with poor risk disease, indicating the higher unmet medical need in these intermediate and poor risk groups. Recently, cabozantinib demonstrated a significant improvement in investigator-assessed PFS and ORR over standard-of-care sunitinib as first-line therapy in a randomized phase 2 study in participants with intermediate- or poor-risk mRCC.

Based on their different mechanisms of action and the potential immune effects associated with anti-VEGF treatment, combining nivolumab with cabozantinib (doublet regimen) or nivolumab and ipilimumab combined with cabozantinib (triplet regimen) may produce synergistic clinical activity and provide improved benefit over standard of care sunitinib monotherapy in mRCC. As mentioned in Section 3.1, several phase 1 studies combining anti-PD-1 agents with anti-VEGF agents have demonstrated greater clinical activity than has been observed historically in trials of these agents given as monotherapy. Although additional risk may be involved with combination therapies, early safety results from the ongoing phase 1 combination study in refractory genitourinary cancers evaluating the safety and efficacy of nivolumab combined with cabozantinib (doublet regimen) and nivolumab and ipilimumab combined with cabozantinib (triplet regimen) suggest that safety profiles appear acceptable. Therefore, overall benefit/risk is acceptable for these doublet and triplet combinations compared to sunitinib in participants with intermediate/poor risk mRCC.

Although the relative benefit of cabozantinib vs sunitinib in participants with previously untreated, favorable risk mRCC was not evaluated in the CABOSUN study, subgroup analyses from the METEOR study suggest that the PFS and OS benefit of cabozantinib in favorable risk disease is similar to that in intermediate risk disease (see Section 3.2.1.4). Based on the preliminary efficacy results from the phase 1 combination study, which included participants with refractory genitourinary tumors, nivolumab combined with cabozantinib (doublet) or nivolumab and ipilimumab combined with cabozantinib (triplet) is anticipated to produce clinical activity that is greater than that observed with sunitinib monotherapy, even in favorable risk disease. Therefore, this study will also include participants with favorable risk disease, to comprise approximately 20% of the overall population.

Overall, the safety profile of nivolumab monotherapy as well as in combination with ipilimumab is manageable and generally consistent across completed and ongoing clinical trials with no

MTD reached at any dose tested up to 10 mg/kg. Most AEs were low-grade (Grade 1 to 2) with relatively few related high-grade (Grade 3 to 4) AEs. There was no pattern in the incidence, severity, or causality of AEs with respect to nivolumab dose level.

A pattern of immune-related adverse events has been defined, for which management algorithms have been developed; these are provided in Appendix 5 and within the Investigator Brochure.[29] Most high-grade events were manageable with the use of corticosteroids or hormone replacement therapy (endocrinopathies) as instructed in these algorithms.

More detailed information about the known and expected benefits and risks and reasonably anticipated adverse events (AEs) of nivolumab, ipilimumab, and cabozantinib may be found in their respective Investigator Brochures.[22,29]

4. OBJECTIVES AND ENDPOINTS

Table 4-1: Objectives and Endpoints

Objective	Endpoint
Primary	
To compare progression-free survival (PFS) per BICR of Arm A (doublet) with Arm C and of Arm B (triplet) with Arm C (sunitinib) in all intermediate/poor risk randomized participants.	The primary objective specifies two comparisons: PFS per BICR of Arm A versus Arm C and of Arm B versus Arm C in all intermediate/poor risk randomized participants. PFS is defined as the time between the date of randomization and the first date of the documented progression, or death due to any cause, whichever occurs first. Participants who die without a reported progression will be considered to have progressed on the date of their death. Participants who did not progress or die will be censored on the date of their last evaluable tumor assessment. Participants who did not have any on study tumor assessments and did not die will be censored on their date of randomization. Participants who started anti-cancer therapy without a prior reported progression will be censored on the date of their last evaluable tumor assessment prior to the initiation of first subsequent anti-cancer therapy.
Secondary	
To compare progression-free survival (PFS) per BICR of Arm A with Arm C and of Arm B with Arm C in all randomized participants.	This secondary objective specifies two comparisons: PFS per BICR of Arm A versus Arm C and of Arm B versus Arm C in all randomized participants. The PFS will be defined similarly to the primary endpoint except that all randomized participants (any risk participants) will be used.
To compare overall survival (OS) of Arm A with Arm C and of Arm B with Arm C in all intermediate/poor risk randomized participants.	The secondary endpoint specifies two comparisons: OS of Arm A versus Arm C and of Arm B versus Arm C in all intermediate/poor risk randomized participants. OS is defined as the time between the date of randomization and the date of death due to any cause. A participant who has not died will be censored at the last known alive date.
To compare overall survival (OS) of Arm A with Arm C and of Arm B with Arm C in all randomized participants.	This secondary objective specifies two comparisons: OS of Arm A versus Arm C and of Arm B versus Arm C in all randomized participants. The OS will be defined similarly to the first secondary endpoint except that all randomized participants (any risk

Table 4-1: Objectives and Endpoints

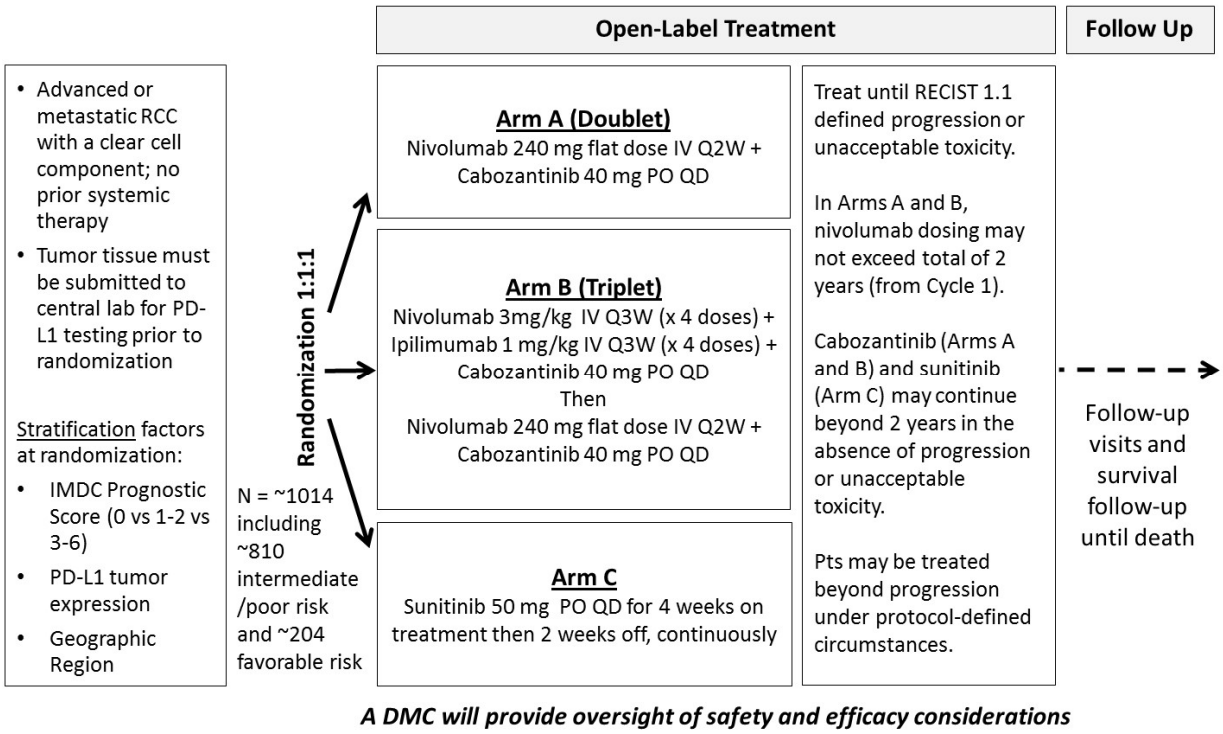
Objective	Endpoint
	participants) will be used.
To evaluate the objective response rate (ORR) per BICR in all intermediate/poor risk randomized and all randomized participants.	This secondary objective is ORR per BICR in all intermediate/poor risk randomized participants and in all randomized participants (any risk participants). ORR is defined as the proportion of randomized participants who achieve a best response of complete response (CR) or partial response (PR) using the RECIST 1.1 criteria. Best overall response (BOR) is defined as the best response designation recorded between the date of randomization and the date of objectively documented progression per RECIST 1.1 or the date of subsequent therapy (including tumor-directed radiotherapy and tumor-directed surgery), whichever occurs first. For participants without document progression or subsequent therapy, all available response designations will contribute to the BOR assessment. Duration of response (DOR) is defined as the time between the date of first confirmed documented response (CR or PR) to the date of first documented tumor progression (per RECIST 1.1) or death due to any cause, whichever occurs first. Participants who neither progress nor die will be censored on the date of their last tumor assessment. Responders who started anti-cancer therapy without a prior reported progression will be censored on the date of their last evaluable tumor assessment prior to the initiation of first subsequent anti-cancer therapy. Time to response (TTR) is defined as the time from randomization to the date of the first confirmed documented response (CR or PR), as assessed by BICR. DOR and TTR will be evaluated for responders (CR or PR) only.
To assess overall safety and tolerability in all treated participants.	As measured by the incidence of adverse events (AEs), serious adverse events (SAEs), AEs leading to discontinuation, deaths, laboratory abnormalities and changes from baseline.
Exploratory	
To explore potential predictive biomarkers of clinical response to nivolumab, ipilimumab and cabozantinib combination.	Analysis of tumor specimens and blood samples for proteins and genes involved in regulating immune response (eg, PD-1, PD-L1, PD-L2, CXCL10, MET). Other exploratory endpoints for biomarkers, pharmacogenomics, and immunogenicity are described in Section 9.8.
To evaluate health related quality of life (HRQoL).	Assessed by the NCCN Functional Assessment of Cancer Therapy-Kidney Symptom Index (FKSI-19) and the EuroQoL Group's EQ-5D (3L version) is described in Section 9.9.
To characterize the pharmacokinetics of nivolumab, ipilimumab, and cabozantinib and explore exposure response relationships, if applicable.	Population PK parameters, exposure-response relationship between select PK measures of exposure and safety and efficacy endpoints, if applicable
To characterize the immunogenicity of nivolumab and ipilimumab.	Incidence of anti-nivolumab and anti-ipilimumab antibodies and their potential relationship with safety and efficacy endpoints

5. STUDY DESIGN

5.1 Overall Design

This is an open label, randomized trial of nivolumab combined with cabozantinib (doublet regimen) or nivolumab + ipilimumab combined with cabozantinib (triplet regimen) versus sunitinib in participants with previously untreated (first line) advanced or metastatic RCC. Participants will be randomized between 3 arms in a 1:1:1 ratio with approximately 810 intermediate/poor risk participants (270 per arm) and 204 favorable risk participants (68 per arm). Participants will be stratified at the time of randomization by IMDC prognostic score (0 [favorable risk] versus 1-2 [intermediate risk] versus 3-6 [poor risk]), PD-L1 tumor expression (≥ 1% versus < 1% or indeterminate), and region (US/Canada/Western Europe/Northern Europe versus rest of the world [ROW]).
The study design schematic is presented in Figure 5.1-1.

Figure 5.1-1: Study Design Schematic



Abbreviations: BICR= blinding independent central review; DMC= data monitoring committee; IMDC= International Metastatic Renal Cell Carcinoma Database Consortium; IV=intravenous; PD= progressive disease; D-L1= programmed death-ligand 1; PO= orally by mouth; Pts=patients/participants; Q2W=every 2 weeks; Q3W=every 3 weeks; QD= once daily; RCC=renal cell carcinoma.

This study will consist of 3 stages: screening, treatment, and follow up phase.

Screening stage: Screening begins by establishing the participant's initial eligibility and signing of the informed consent (ICF). Sufficient, recent tumor tissue, preferably obtained within 3 months but no more than 12 months prior to enrollment, from a metastatic tumor lesion or from a primary tumor lesion which has not been previously irradiated (formalin-fixed paraffin-embedded block or 20 unstained slides: a minimum of 10 slides, obtained from core biopsy,

punch biopsy, excisional biopsy or surgical specimen) will be submitted to the central laboratory. Upon receipt of the tumor sample, the central lab will determine PD-L1 expression level by immunohistochemistry (IHC) testing (see Section 9.8.2). In order to be randomized in the Interactive Response Technology (IRT) system, a participant must be classified as PD-L1 expression ≥ 1%, PD-L1 expression < 1%, or PD-L1 expression indeterminate. Sites will be informed when the submitted tumor sample is insufficient for PD-L1 testing by the central lab.

Participants will be assessed for complete study eligibility prior to randomization as specified in Section 2.

The Screening stage ends with either confirmation of full eligibility and randomization for the participant or with the confirmation that the participant is a screen failure. This study permits the re-enrollment of a participant that has discontinued the study as a pre-treatment failure prior to randomization. If re-enrolled, the participant must be re-consented. A new participant identification number will be assigned by IRT at the time of re-enrollment.

Treatment stage: The Treatment stage begins when the randomization call is made into the IRT. The participant is randomly assigned to 1 of the 3 treatment arms as noted in the study schematic above.

- Arm A (Doublet): Nivolumab 240 mg flat dose IV Q2W + Cabozantinib 40 mg PO QD
  - Nivolumab to be continued until disease progression, unacceptable toxicity, or a maximum of 2 years from the first dose in Cycle 1
  - Cabozantinib to be continued until disease progression or unacceptable toxicity
- Arm B (Triplet): Nivolumab 3mg/kg IV + Ipilimumab 1 mg/kg IV, both Q3W x 4 doses + Cabozantinib 40 mg PO QD
  - Then Nivolumab 240 mg flat dose IV Q2W + Cabozantinib 40 mg PO QD
  - Nivolumab to be continued until disease progression, unacceptable toxicity, or a maximum of 2 years from the first dose in Cycle 1
  - Cabozantinib to be continued until progression or unacceptable toxicity
- Arm C: Sunitinib 50 mg PO QD for 4 weeks, followed by 2 weeks off, per cycle. Cycles to be continued until progression or unacceptable toxicity

Study treatment must begin within 3 days of randomization. Participants in Arm A and Arm B will continue nivolumab until progression, unacceptable toxicity, withdrawal of consent, or a maximum of 2 years from the first dose in Cycle 1, whichever occurs first. Cabozantinib (Arm A and Arm B) may be continued until progression, unacceptable toxicity, or withdrawal of consent, whichever occurs first, and may extend beyond 2 years from the first dose in Cycle 1. See Table 7.1-1 and Table 7.1-2 for the dosing schedule. Study drugs may be delayed for toxicity (See Section 7.4.1). Treatment may be continued beyond investigator-assessed progression if the investigator confirms that the participant meets the criteria specified in Section 8.1.4 and approval is granted by the BMS Medical Monitor.

A negative pregnancy test should be documented within 24 hours prior to the initial dose of the investigational product and then performed every 4 weeks ± 7 days during treatment. On-study

laboratory assessments should be drawn within 72 hours prior to dosing and will be assessed at the local laboratory.

Tumor assessments will occur in accordance with Section 2 and Section 9.1.2 until progression has been assessed by the investigator **and** confirmed by the blinded independent central review (BICR). Each site must submit scans on a rolling basis, preferably within 7 days of image acquisition, to a third-party vendor for BICR. If progression is assessed by the investigator, the site will inform the radiology vendor so that the BICR assessment of progression can be performed. The BICR assessment of progression will be completed, and the results provided to the site, within approximately 14 days, as specified in Section 9.1.2.

PK and immunogenicity samples will be collected according to Table 9.5-1 and Table 9.5-2. Adverse event assessments should be documented at each clinic visit.

Quality of Life will be assessed using the National Comprehensive Cancer Network (NCCN) Functional Assessment of Cancer Therapy - Kidney Symptom Index (FKSI-19) and EuroQoL's EQ-5D-3L. These questionnaires should be completed according to Section 2.

The Treatment stage ends when the participant is discontinued from study therapy.

Follow-up stage: The Follow-up stage begins when the decision to discontinue a participant from study therapy is made (no further treatment with study therapy). Participants will have 2 follow-up visits (FU1 and FU2) for safety within approximately 100 days from the last dose of study therapy, and AEs will be followed until the toxicities resolve, return to baseline, or are deemed irreversible. The FKSI-19 and EQ-5D-3L will be completed as described in Section 2.

Participants who discontinued study treatment without BICR confirmed radiographic progression will continue to have tumor assessments performed according to the frequency described in Sections 2 and 9.1.2, even if new anti-tumor therapy has been initiated. If progression is assessed by the investigator, the site will inform the radiology vendor, so that the BICR assessment of progression can be performed. Participants whose progression is not confirmed by the BICR will be required to continue tumor assessments (if clinically feasible) according to the protocol-specified schedule or sooner if clinically indicated until BICR confirms progression on a subsequent tumor assessment (See Section 9.1.2.2 for additional details).

After the Follow-up 2 Visit, all participants will be followed for overall survival status every 3 months (+/- 14 days) until death, withdrawal of consent, loss to follow-up, or end of study. Survival status can be ascertained in person or by telephone contact. If new anti-tumor therapy is initiated for either progression or a secondary malignancy at any time during this period, this and all other pertinent data obtained should be recorded on the appropriate Case Report Form (CRF).

5.1.1 Data Monitoring Committee and Other External Committees

To provide independent oversight of safety, efficacy, and study conduct, a data monitoring committee (DMC) will be instituted. The DMC will meet regularly to ensure that participant safety is carefully monitored, including a safety assessment after the first 75 participants are randomized across the 3 arms and are followed for at least 6 weeks. The DMC will convene additional ad hoc meetings if necessary. Following each meeting, the DMC will recommend

continuation, modification, or discontinuation of the study based on observed toxicities. The DMC will also review the interim analysis results and inform BMS whether stopping criteria for superiority are met at that time. A separate DMC charter will describe the activities of this committee in more detail.

Blinded independent central review (BICR) will be utilized in this study for determination of PFS and ORR endpoints. The BICR will review all available tumor assessment scans for all randomized participants. Details of BICR responsibilities and procedures will be specified in the BICR charter.

5.2 Number of Participants

Approximately 1353 participant will be enrolled in order to randomize approximately 1014 participants (338 per arm). This includes approximately 204 favorable risk (per IMDC criteria) randomized participants (68 per arm) and approximately 810 intermediate/poor risk (per IMDC criteria) randomized participants (270 per each arm). The number of randomized participants with favorable risk disease will be capped at approximately 204 participants.

5.3 End of Study Definition

The start of the trial is defined as the first visit for the first participant screened. End of trial is defined as the last visit or scheduled procedure shown in the Schedule of Activities for the last participant. Study completion is defined as the final date on which data for the endpoint of overall survival was or is expected to be collected, if this is not the same.

5.4 Scientific Rationale for Study Design

Please refer to Section 3.1 for details of the rationale for evaluating the combination of nivolumab with cabozantinib (doublet regimen) or nivolumab and ipilimumab combined cabozantinib (triplet regimen) in this study.

5.4.1 Rationale for Open-Label Design

The trial will have an open-label design with a blinded independent central review. Different dosing schedules and use of both IV infusions and oral medications across the 3 different arms make blinding the trial impractical. The complexity of including multiple visits for placebo infusions are burdensome for this participant population. Nivolumab and ipilimumab are associated with some toxicities that are also common to cabozantinib and sunitinib (eg, diarrhea, hepatotoxicity), and the management of these common toxicities are different for immuno-oncology agents (eg, typically requiring steroids) vs TKIs (eg, typically requiring dose delay and reductions). Therefore, an open-label trial is preferable for participant safety as it allows the optimal management of toxicities.

5.4.2 Rationale for Choice of Primary Endpoint in Intermediate and Poor-Risk Participants

The primary endpoint to be evaluated is PFS per BICR in participants with intermediate or poor risk mRCC according to IMDC prognostic criteria. PFS has been an acceptable endpoint to support regulatory approvals in first-line mRCC. Sorafenib, sunitinib, bevacizumab, and

pazopanib were approved based on an increase in PFS. While OS remains the "gold standard" for oncology trials, as more active agents gain approval in RCC, subsequent treatment with these agents confounds the effect of first-line treatment on OS. Therefore, PFS remains an important indicator of clinical benefit in first-line mRCC and has been chosen as the primary endpoint for this trial.

The primary efficacy analysis population is limited to intermediate and poor-risk participants, which comprises approximately 75% of the total treatment-naive mRCC population. The outcomes associated with standard of care first-line treatment for participants with intermediate and poor risk disease is suboptimal, and selecting these participants for the primary endpoint of the study will allow for potential meaningful differences in efficacy to be detected earlier than in a broader population that includes favorable-risk participants. Nevertheless, this trial will also randomize approximately 204 favorable-risk participants in order to evaluate the clinical activity and safety in participants across all risk groups (ie, all randomized population), which are included as secondary objectives. Based on results for favorable-risk subjects in the METEOR trial, efficacy would be expected in favorable-risk subjects in the current trial?

5.5 Justification for Dose

5.5.1 Dosing for Sunitinib Monotherapy in Arm C

Dosing for sunitinib monotherapy in Arm C uses the standard dosing schedule found in the sunitinib prescribing information.[34]

5.5.2 Dosing for Cabozantinib Doublet and Triplet Therapy in Arm A and Arm B

Dosing for nivolumab combined with cabozantinib (doublet regimen) or nivolumab and ipilimumab combined with cabozantinib (triplet regimen) was determined based on a phase 1b/2 trial (NCT02496208) conducted with the combination at NCI/NIH (see Section 3.2.1.5).

5.5.3 Dosing Rationale for Nivolumab 240 mg flat dose Q2W

A flat dose of nivolumab 240 mg every 2 weeks will be administered in Arm A and in the maintenance phase of Arm B in combination with cabozantinib. Nivolumab 3mg/kg once every 2 weeks has been shown to be similar to 240 mg flat dose every 2 weeks, and is currently FDA approved, based on clinical data and modeling and simulation approaches using population PK (PPK) and exposure-response analyses of data from studies in multiple tumor types (melanoma, non-small-cell lung cancer [NSCLC], renal cell carcinoma [RCC]) and urothelial carcinoma where body weight normalized dosing (mg/kg) has been used.

PPK analyses have shown that the PK of nivolumab is linear with proportional exposure over a dose range of 0.1 to 10 mg/kg, and no differences in PK across ethnicities and tumor types were observed. Nivolumab clearance and volume of distribution were found to increase as the body weight increases, but less than proportionally with increasing weight, indicating that mg/kg dosing represents an over-adjustment for the effect of body weight on nivolumab PK. The PPK model developed using data from NSCLC participants was updated, using data from participants investigating nivolumab in the treatment of melanoma, NSCLC, and RCC. In this dataset, the

median (minimum to maximum) weight was 77 kg (35 to 160 kg) and thus, an approximately equivalent dose of 3 mg/kg for an 80 kg participant, nivolumab 240 mg Q2W was selected for future studies. To predict relevant summary exposures of nivolumab 240 mg Q2W, the PPK model was used to simulate nivolumab 3 mg/kg Q2W and 240 mg Q2W. In the simulations, the simulated participant populations consisted of 1000 participants per treatment arm randomly sampled from aforementioned pooled database of cancer participants. Because no differences in PK were noted across ethnicities and tumor types, these simulated melanoma, NSCLC and RCC data will be applicable to participants with other tumor types. The simulated measure of exposure of interest, time-averaged concentrations (Cavgss) for 240 mg Q2W were predicted to be similar for all participants in reference to 80 kg participants receiving 3 mg/kg Q2W.

Additionally, nivolumab is safe and well tolerated up to 10 mg/kg Q2W dose level. Adverse events have been consistent across tumor types following monotherapy and have not demonstrated clear dose-response or exposure-response relationships. Additionally, the simulated median and 95th prediction interval of nivolumab summary exposures across a wide body weight range (35 - 160 kg) are predicted to be maintained below the corresponding observed highest exposure experienced in nivolumab, ie, 95th percentile following nivolumab 10 mg/kg Q2W from clinical study CA209003. Thus, while participants in the lower body weight ranges would have greater exposures than 80 kg participants, the exposures are predicted to be within the range of observed exposures at doses (up to 10 mg/kg Q2W) used in the nivolumab clinical program, and are not considered to put participants at increased risk. Additionally, for participants with greater body weights, the simulated ranges of exposures are also not expected to affect efficacy, because the exposures predicted following administration of a 240 mg Q2W are on the flat part of the exposure-response curves for previously investigated tumors, melanoma and NSCLC. Given the similarity of nivolumab PK across many tumor types, including RCC, and the similar exposures predicted following administration of 240 mg flat dose compared to 3 mg/kg, it was shown that the safety and efficacy profile of 240 mg nivolumab will be similar to that of 3 mg/kg nivolumab.

5.5.4 Rationale for Dosing Duration for Nivolumab

The optimal duration of immunotherapy is currently unknown. However, because immunotherapy engages the immune system to control the tumor, continuous treatment as is required with targeted agents or cytotoxic therapy may not be necessary.

Accumulating evidence from different clinical trials in different tumor types with nivolumab or nivolumab combined to ipilimumab indicates that most of the responses are generally occurring early, with a median time to response of 2 to 4 months.[46,47] A recent analysis in a melanoma study suggests the majority of participants who discontinue nivolumab and/or ipilimumab for toxicity maintain disease control in the absence of further treatment.[48] Furthermore, a limited duration of ipilimumab including only 4 induction doses resulted in long term survival in participants with metastatic melanoma, with a sustained plateau in survival starting at around year 3.[49]

For these reasons, in study CA2099ER, treatment with nivolumab or nivolumab with ipilimumab will be given for up to 24 months in the absence of disease progression or unacceptable toxicity. Participants who complete 24 months of nivolumab or nivolumab with ipilimumab and have subsequent disease progression may reinitiate nivolumab or nivolumab with ipilimumab at the same dose and schedule given previously on study and continue such treatment for up to 1 additional year.

6. STUDY POPULATION

For entry into the study, the following criteria MUST be met.

6.1 Inclusion Criteria

1) Signed Written Informed Consent

a) Participants must have signed and dated an IRB/IEC approved written informed consent form in accordance with regulatory and institutional guidelines. This must be obtained before the performance of any protocol related procedures that are not part of normal participant care.

b) Participants must be willing and able to comply with scheduled visits, treatment schedule, and laboratory testing.

2) Type of Participant and Target Disease Characteristics

a) Histological confirmation of RCC with a clear-cell component, including participants who may also have sarcomatoid features

b) Advanced (not amenable to curative surgery or radiation therapy) or metastatic (AJCC Stage IV) RCC

c) No prior systemic therapy for RCC with the following exception:

i) One prior adjuvant or neoadjuvant therapy for completely resectable RCC if such therapy did not include an agent that targets VEGF or VEGF receptors and if recurrence occurred at least 6 months after the last dose of adjuvant or neoadjuvant therapy.

d) Karnofsky Performance Status (KPS) ≥ 70% (See Appendix 7)

e) Measurable disease as per RECIST v1.1 per investigator. (See Appendix 8)

f) Either a formalin-fixed, paraffin-embedded (FFPE) tissue block or unstained tumor tissue sections, preferably obtained within 3 months but no more than 12 months prior to enrollment, with an associated pathology report, must be submitted to the central laboratory during screening. Biopsy should be excisional, incisional or core needle. Fine needle aspiration is unacceptable for submission. In participants who received prior adjuvant or neoadjuvant therapy, the tumor sample must have been obtained after completion of adjuvant or neoadjuvant therapy. Upon receipt of the tumor sample, the central lab will determine PD-L1 expression level by IHC testing (see Section 9.8.2). In order to be randomized in the IRT system, a participant must be classified as PD-L1 expression ≥ 1%, PD-L1 expression < 1%, or PD-L1 expression indeterminate. Sites will be informed when the submitted tumor sample is insufficient for PD-L1 testing by the central lab.

g) Participants with favorable, intermediate and poor risk categories will be eligible for the study. To be eligible for the Intermediate and Poor-Risk cohort, at least one of the

following prognostic factors as per International Metastatic RCC Database Consortium (IMDC) must be present: (See Appendix 6)

i) KPS equal to 70%

ii) Less than 1 year from initial diagnosis (including original localized disease if applicable) to randomization

iii) Hemoglobin < lower limit of normal (LLN)

iv) Corrected calcium concentration > 10 mg/dL

v) Absolute neutrophil count > ULN

vi) Platelet count > ULN

If none of the above factors are present, participants are only eligible for the favorable-risk cohort. Approximately 204 favorable risk participants will be randomized. Enrollment of favorable risk participants may be closed earlier than enrollment for intermediate and poor risk participants.

3) Age and Reproductive Status

a) Males and Females, ages 18 or age of majority

b) Women of childbearing potential (WOCBP) must have a negative serum or urine pregnancy test (minimum sensitivity 25 IU/L or equivalent units of HCG) within 24 hours prior to the start of study treatment

c) Women must not be breastfeeding

d) Women of childbearing potential (WOCBP) must agree to follow instructions for method(s) of contraception for the duration of study treatment and 5 months after the last dose of study treatment (ie, 30 days [duration of ovulatory cycle] plus the time required for the investigational drug to undergo approximately five half-lives).

e) Males who are sexually active with WOCBP must agree to follow instructions for method(s) of contraception for the duration of study treatment and 7 months after the last dose of study treatment (ie, 90 days [duration of sperm turnover] plus the time required for the investigational drug to undergo approximately five half-lives).

f) Azoospermic males are exempt from contraceptive requirements. WOCBP who are continuously not heterosexually active are also exempt from contraceptive requirements, and still must undergo pregnancy testing as described in this section.

Investigators shall counsel WOCBP, and male participants who are sexually active with WOCBP, on the importance of pregnancy prevention and the implications of an unexpected pregnancy. Investigators shall advise on the use of highly effective methods of contraception (Appendix 4) which have a failure rate of < 1% when used consistently and correctly.

6.2 Exclusion Criteria

1) Medical Conditions

a) Any active CNS metastases. Participants with treated, stable CNS metastases for at least 3 months are eligible as long as they meet the following criteria:

Treated CNS metastases are defined as having no ongoing requirement for corticosteroids for at least 2 weeks prior to randomization and no evidence of progression or hemorrhage after treatment completed at least 3 months prior to randomization, as ascertained by

clinical examination and brain imaging (MRI or CT). (Stable dose of anticonvulsants is allowed). Treatment for CNS metastases may include whole brain radiotherapy, radiosurgery (eg, RS, gamma knife, LINAC, or equivalent) or a combination as deemed appropriate by the treating physician. Participants with CNS metastases treated by neurosurgical resection or brain biopsy performed within 3 months prior to randomization are not eligible. Baseline imaging of the brain is required within 28 days prior to randomization.

b) Any active, known or suspected autoimmune disease. Participants with type I diabetes mellitus, hypothyroidism only requiring hormone replacement, skin disorders (such as vitiligo, psoriasis, or alopecia) not requiring systemic treatment, or conditions not expected to recur in the absence of an external trigger are permitted to enroll.

c) Any condition requiring systemic treatment with either corticosteroids (> 10 mg daily prednisone equivalent) or other immunosuppressive medications within 14 days of randomization. Inhaled or topical steroids, and adrenal replacement steroid doses > 10 mg daily prednisone equivalent, are permitted in the absence of active autoimmune disease.

d) Prior malignancy active within the previous 3 years except for locally curable cancers that have been apparently cured, such as basal or squamous cell skin cancer, superficial bladder cancer, or carcinoma in situ of the prostate, cervix, or breast

e) Any tumor invading the SVC or other major blood vessels

f) Any tumor invading the GI tract or any evidence of endotracheal or endobronchial tumor within 30 days prior to randomization

g) Known history of positive test for human immunodeficiency virus (HIV) or known acquired immunodeficiency syndrome (AIDS). NOTE: Testing for HIV must be performed at sites where mandated locally (sites in Germany, see Appendix 12).

h) Known medical condition (eg, a condition associated with diarrhea or acute diverticulitis, aortic aneurysm, aortic dissection) that, in the investigator's opinion, would increase the risk associated with study participation or study drug administration or interfere with the interpretation of safety results

i) History of abdominal fistula, gastrointestinal perforation, intra-abdominal abscess, bowel obstruction, or gastric outlet obstruction within the past 6 months prior to randomization

j) Impairment of gastrointestinal function or gastrointestinal disease that may significantly alter the absorption of cabozantinib or sunitinib (eg, malabsorptive disorder, ulcerative disease, uncontrolled nausea, vomiting, diarrhea, or small bowel resection)

k) Serious, non-healing wound or ulcer within 30 days prior to randomization

l) Evidence of active bleeding or bleeding susceptibility; or medically significant hemorrhage within prior 3 months prior to randomization

m) Uncontrolled adrenal insufficiency

n) History of cerebrovascular accident (CVA) including transient ischemic attack within the past 6 months prior to randomization

o) History of deep vein thrombosis (DVT) or pulmonary embolism (PE) within past 6 months prior to randomization unless stable, asymptomatic, and treated with low molecular weight heparin (LMWH) for at least 6 weeks prior to randomization

p) Any unstable cardiac arrhythmia within 6 months prior to randomization

q) Prolongation of the Fridericia corrected QT (QTcF) interval defined as > 450 msec for males and > 470 msec for females, where QTcF = QT / $3\sqrt{RR}$ with triplicate measurements

r) Poorly controlled hypertension (defined as systolic blood pressure (SBP) of > 150 mmHg or diastolic blood pressure (DBP) of > 90 mmHg), despite antihypertensive therapy

s) History of any of the following cardiovascular conditions within 6 months of randomization: cardiac angioplasty or stenting, myocardial infarction, unstable angina, coronary artery by-pass graft surgery, symptomatic peripheral vascular disease, class III or IV congestive heart failure (CHF), as defined by the New York Heart Association (NYHA)

t) Any radiologic or clinical evidence of pancreatitis within 30 days prior to randomization

u) Inability to swallow oral medications

2) Prior/Concomitant Therapy

a) Prior treatment with VEGF, MET, AXL, KIT, or RET targeted therapy (including, but not limited to, sunitinib, pazopanib, axitinib, tivozanib, sorafenib, lenvatinib, bevacizumab, and cabozantinib).

b) Prior treatment with an anti-PD-1, anti-PD-L1, anti-PD-L2, anti-CD137, or anti-CTLA-4 antibody, or any other antibody or drug specifically targeting T-cell co-stimulation or checkpoint pathways.

c) Concomitant strong CYP3A4 inducers or inhibitors within 14 days prior to randomization. Please refer to Appendix 9 for a list of common strong CYP3A4 inducers and inhibitors.

d) Concomitant treatment, in therapeutic doses, with anticoagulants such as warfarin or warfarin-related agents, thrombin or Factor Xa inhibitors. Aspirin (up to 325 mg/day) and prophylactic and therapeutic low molecular weight heparin (LMWH) are permitted.

e) Major surgery (eg, nephrectomy) less than 6 weeks prior to randomization

f) Any of the following prior radiotherapy procedures:

i) Radiotherapy to the thoracic cavity or abdomen within 4 weeks prior to randomization

ii) Radiotherapy to bone lesions within 2 weeks prior to randomization

iii) Radiotherapy to any other site within 4 weeks prior to randomization

NOTE: In all cases, there must be complete recovery and no ongoing complications from prior radiotherapy.

3) Physical and Laboratory Test Findings

a) Ejection fraction ≤ 50% on screening echocardiogram or MUGA

b) WBC < 2000/μL

c) Neutrophils < 1500/μL

d) Platelets < 100 x 10^3/μL

e) Hemoglobin < 9.0 g/dL (support with transfusion is acceptable)

f) Serum creatinine > 1.5 x ULN or calculated creatinine clearance < 40 mL/min (using the Cockcroft-Gault formula)

g) AST/ALT > 3.0 x ULN

h) Total bilirubin > 1.5 x ULN (except participants with Gilbert Syndrome who must have a total bilirubin level of < 3.0 x ULN)

i) Urine protein/creatinine ratio (UPCR) ≥ 1.0, unless 24-hour urine protein is < 1.0 g

j) INR > 1.2

k) Any positive test result for hepatitis B virus or hepatitis C virus indicating presence of virus, eg, hepatitis B surface antigen (HBsAg) positive, or hepatitis C antibody (anti-HCV) positive (except if HCV-RNA negative)

4) Allergies and Adverse Drug Reaction

a) History of allergy or hypersensitivity to study drug components

b) No history of severe hypersensitivity to a monoclonal antibody

5) Other Exclusion Criteria

a) Prisoners or participants who are involuntarily incarcerated. (Note: under certain specific circumstances a person who has been imprisoned may be included or permitted to continue as a participant. Strict conditions apply and Bristol-Myers Squibb approval is required.

b) Participants who are compulsorily detained for treatment of either a psychiatric or physical (eg, infectious disease) illness

Eligibility criteria for this study have been carefully considered to ensure the safety of the study participants and that the results of the study can be used. It is imperative that participants fully meet all eligibility criteria.

6.3 Lifestyle Restrictions

6.3.1 Meals and Dietary Restrictions

For participants in Arms A and B, cabozantinib should not be taken with grapefruit/grapefruit juice or Seville oranges.

For participants in Arm C, sunitinib should not be taken with grapefruit/grapefruit juice.

6.4 Screen Failures

Screen failures are defined as participants who consent to participate in the clinical study but are not subsequently randomized. A minimal set of screen failure information is required to ensure transparent reporting of screen failure participants, to meet the Consolidated Standards of Reporting Trials (CONSORT) publishing requirements, and to respond to queries from regulatory authorities. Minimal information includes date of consent, demography, screen failure details, eligibility criteria, and any serious AEs.

6.4.1 Retesting During Screening or Lead-In Period

Participant Re-enrollment: This study permits the re-enrollment of a participant that has discontinued the study as a pre-treatment failure (ie, participant has not been randomized / has not been treated). If re-enrolled, the participant must be re-consented.

Retesting of laboratory parameters and/or other assessments within any single Screening or Lead-in period will be permitted (in addition to any parameters that require a confirmatory value).

The most current result prior to randomization is the value by which study inclusion will be assessed, as it represents the participant's most current, clinical state.

Laboratory parameters and/or assessments that are included in Table 2-1 Screening Procedural Outline may be repeated in an effort to find all possible well-qualified participants. Consultation with the BMS Medical Monitor may be needed to identify whether repeat testing of any particular parameter is clinically relevant.

7. TREATMENT

Study treatment is defined as any investigational treatment(s), marketed product(s), placebo or medical device intended to be administered to a study participant according to the study randomization or treatment allocation.

Study treatment includes both Investigational [Medicinal] Product (IP/IMP) and Non-investigational [Medicinal] Product (Non-IP/Non-IMP) and can consist of the following:

An investigational product, also known as investigational medicinal product in some regions, is defined a pharmaceutical form of an active substance or placebo being tested or used as a reference in a clinical study, including products already with a marketing authorization but used or assembled (formulated or packaged) differently than the authorized form, or used for an unauthorized indication, or when used to gain further information about the authorized form.

Other medications used as support or escape medication for preventative, diagnostic, or therapeutic reasons, as components of the standard of care for a given diagnosis, may be considered as non-investigational products.

Table 7-1: Study treatments for CA2099ER

Product Description / Class and Dosage Form	Potency	IP/ Non-IMP	Blinded or Open Label	Packaging / Appearance	Storage Conditions (per label)
BMS-936558-01 (Nivolumab) Solution for Injection*	100 mg (10 mg/mL)	IP	Open Label	Vial (multiple vials per carton)	Store at 2° - 8 °C. Protect from light and freezing
Ipilimumab Solution for Injection	200 mg (5mg/mL)	IP	Open Label	Vial (multiple vials per carton)	Store at 2° - 8 °C. Protect from light and freezing
Cabozantinib Tablet	20 mg	IP	Open Label	Tablets in a bottle	Refer to storage conditions on container label
Sunitinib Malate Capsule**	12.5 mg	IP	Open Label	Capsules in various packaging configurations	Refer to storage conditions on container label/package insert

*May be labeled as "BMS-936558-01" or "Nivolumab"

** Sunitinib may be obtained by the investigational sites in certain countries as local commercial product (which may be available as a different potency/package size than listed above) if local regulations allow and agreed to by BMS.

Premedications or medications used to treat infusion-related reactions should be sourced by the investigative sites if available and permitted by local regulations. Solutions used as diluent (ie, 0.9% Sodium Chloride Injection or 5% Dextrose Injection) should also be sourced by investigative sites if available and permitted by local regulations.

7.1 Treatments Administered

The selection and timing of dose for each participant is presented in Table 7.1-1 and the dosing scheduled by cycle is presented in Table 7.1-2.

Table 7.1-1: Selection and Timing of Dose

Arm	Study Treatment	Dosage level(s) and Formulation	Frequency of Administration	Route of Administration
A (Doublet)	Nivolumab	240 mg IV	Every 2 weeks	IV
	Cabozantinib	40 mg (20 mg tablets)	Once daily (QD)	PO
B (Triplet)	Nivolumab	3 mg/kg IV for 4 doses then 240 mg IV	Every 3 weeks (Q3W) for 4 doses then every 2 weeks (Q2W)	IV
	Ipilimumab	1 mg/kg IV for 4 doses	Every 3 weeks (Q3W) for 4 doses	IV
	Cabozantinib	40 mg (20 mg tablets)	Once daily (QD)	PO
C	Sunitinib	50 mg (12.5 mg capsules)	Every 3 weeks (Q3W) for 4 doses then every 2 weeks (Q2W)	IV

Abbreviations: IV=intravenous; PO= by mouth

Table 7.1-2: Dosing Schedule for CA2099ER

Arm	Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5 and subsequent cycles
Arm A Nivolumab + Cabozantinib (each cycle = 2 weeks)	Nivolumab 240 mg IV on Day 1 + Cabozantinib 40 mg PO QD	Nivolumab 240 mg IV on Day 1 + Cabozantinib 40 mg PO QD	Nivolumab 240 mg IV on Day 1 + Cabozantinib 40 mg PO QD	Nivolumab 240 mg IV on Day 1 + Cabozantinib 40 mg PO QD	Nivolumab 240 mg IV on Day 1 + Cabozantinib 40 mg PO QD
Arm B Nivolumab + Ipilimumab + Cabozantinib (Cycle 1 to 4 = 3 weeks, Cycle 5 and subsequent cyces = 2 weeks)	Nivolumab 3 mg/kg IV on Day 1 + Ipilimumab 1 mg/kg IV on Day 1 + Cabozantinib 40 mg PO QD	Nivolumab 3 mg/kg IV on Day 1 + Ipilimumab 1 mg/kg IV on Day 1 + Cabozantinib 40 mg PO QD	Nivolumab 3 mg/kg IV on Day 1 + Ipilimumab 1 mg/kg IV on Day 1 + Cabozantinib 40 mg PO QD	Nivolumab 3 mg/kg IV on Day 1 + Ipilimumab 1 mg/kg IV on Day 1 + Cabozantinib 40 mg PO QD	Nivolumab 240 mg IV on Day 1 + Cabozantinib 40 mg PO QD
Arm C Sunitinib (each cycle = 6 weeks, treatment for 4 weeks then 2 weeks off each cycle)	Sunitinib 50 mg PO QD x 4 weeks	Sunitinib 50 mg PO QD x 4 weeks	Sunitinib 50 mg PO QD x 4 weeks	Sunitinib 50 mg PO QD x 4 weeks	Sunitinib 50 mg PO QD x 4 weeks

Abbreviations: IV=intravenous; PO= by mouth; QD= once daily

Participants should begin study treatment within 3 calendar days of randomization.

For **Arm A**, participants should receive nivolumab at a dose of 240 mg as an approximately 30 minute infusion on Day 1 of each treatment cycle until progression, unacceptable toxicity, withdrawal of consent, completion of 24 months of treatment (from the first dose on Cycle 1), or the end of the study, whichever occurs first. The first cabozantinib dose should be given in the evening on Day 1, Cycle 1 (after the Cycle 1 nivolumab dose).

For **Arm B**, when nivolumab and ipilimumab are to be administered on the same day, nivolumab is to be administered first. Nivolumab infusion (approximately 30 minutes) must be promptly followed by a saline flush to clear the line of nivolumab before starting the ipilimumab infusion. The second infusion will always be the ipilimumab study drug (approximately 30 minutes infusion) and will start after the infusion line has been flushed, filters changed and participant has been observed to ensure no infusion reaction has occurred. The time in between infusions is expected to be approximately 30 minutes (from the end of the nivolumab infusion to the start of the ipilimumab infusion) but may be more or less depending on the situation. The first cabozantinib dose should be given in the evening on Day 1, Cycle 1 (after the Cycle 1 nivolumab and ipilimumab doses).

Starting with Cycle 5 in Arm B, participants should receive nivolumab at a dose of 240 mg as a 30 minute infusion on Day 1 of each treatment cycle until progression, unacceptable toxicity, withdrawal of consent, completion of 24 months of treatment (from the first dose on Cycle 1), or the end of the study, whichever occurs first.

Dosing calculations for nivolumab and ipilimumab should be based on the body weight assessed at baseline. It is not necessary to re-calculate subsequent doses if the participant weight is within 10% of the weight used to calculate the previous dose. All doses should be rounded up or to the nearest milligram per institutional standard.

There will be no dose escalations or reductions of nivolumab or ipilimumab allowed. Participants may be dosed with nivolumab no less than 12 days from the previous dose during q2w cycles. Participants may be dosed with nivolumab and ipilimumab no less than 19 days from the previous dose during q3w cycles (ie, Cycles 1-4 in Arm B). Participants in Arm B may be dosed with the first nivolumab maintenance dose (Cycle 5) no less than 19 days from the previous nivolumab and ipilimumab doses. Premedications are not recommended for the first dose of nivolumab or ipilimumab.

Participants should be carefully monitored for infusion reactions during nivolumab and/or ipilimumab administration. If an acute infusion reaction is noted, participants should be managed according to Section 7.4.4.

Doses of nivolumab and/or ipilimumab may be interrupted, delayed, or discontinued depending on how well the participant tolerates the treatment. Dosing visits are not skipped, only delayed. The only exception is Cycle 4 in Arm B, which may be skipped (omitted) only for the reasons specified in Section 7.4.3.1and Section 8.1.1 below.

7.1.1 Nivolumab

Nivolumab Injection, 100 mg/10 mL (10 mg/mL) and 40 mg/mL (10 mg/mL) Nivolumab injection is to be administered as an IV infusion through a 0.2-micron to 1.2-micron pore size, low-protein binding in-line filter at the protocol-specified doses. It is not to be administered as an IV push or bolus injection. Please refer to the Investigational Brochure/pharmacy manual for further details regarding storage, preparation and administration. Care must be taken to assure sterility of the prepared solution as the product does not contain any antimicrobial preservative or bacteriostatic agent.

7.1.2 Ipilimumab

For details regarding ipilimumab storage, preparation, and administration, please refer to the instructions in the ipilimumab IB and/or pharmacy manual.

Separate infusion bags and filters should be used when administering nivolumab and ipilimumab on the same day.

Care must be taken to assure sterility of the prepared solutions, since the drug product does not contain any antimicrobial preservatives or bacteriostatic agents.

7.1.3 Cabozantinib

Cabozantinib is taken orally on an empty stomach, preferably at bed time. Participants should fast (with the exception of water) for at least 2 hours before until 1 hour after each dose of cabozantinib. Cabozantinib tablets should not be crushed or chewed. Cabozantinib should not be taken with grapefruit/grapefruit juice or Seville oranges. Missed doses of cabozantinib should not be taken within 12 hours of the next dose.

Dispense cabozantinib tablets in their original containers.

7.1.4 Sunitinib

Sunitinib is taken orally without regard to meals. Participants are to avoid grapefruit/grapefruit juice while on treatment with sunitinib.

7.2 Method of Treatment Assignment

All participants will be centrally randomized using an Interactive Response Technology (IRT). Before the study is initiated, each user will receive log in information and directions on how to access the IRT. Specific instructions for using IRT will be provided to the investigational site in a separate document.

The investigator or designee will register the participant for enrollment by following the enrollment procedures established by BMS.

Once enrolled in IRT, enrolled participants that have met all eligibility criteria, including determination of PD-L1 expression in the tumor sample submitted to and evaluated by the central laboratory, will be ready to be randomized through the IRT prior to the start of study treatment administration for each participant. The site will record the treatment assignment on the applicable case report form, if required.

Study treatment will be dispensed at the study visits as listed in Schedule of Activities (Section 2).

7.3 Blinding

Not applicable as this is an open-label study; however, the specific treatment to be taken by a participant will be assigned using an Interactive Response Technology (IRT).

Treatment assignments will be released to the bioanalytical laboratory in order to minimize unnecessary analysis of samples.

BICR will remain blinded to treatment assignment.

7.4 Dosage Modification

In Arm A and Arm B, the assessment of causality for any AE should be performed independently for each study drug in the combination regimen.

In the doublet treatment arm (Arm A), cabozantinib and nivolumab will be administered at Cycle 1. After Cycle 1, modifications in cabozantinib dosing (delay, reduction/escalation, and discontinuation) may occur as outlined in Section 7.4.1, Section 7.4.2, and Section 8. After Cycle 1, nivolumab may be modified as outlined in Section 7.4.1, Section 7.4.2, and Section 8.

In the triplet treatment arm (Arm B), cabozantinib, nivolumab, and ipilimumab will be administered at Cycle 1. After Cycle 1, modifications in cabozantinib dosing (delay, reduction/escalation, and discontinuation) may occur as outlined in Section 7.4.1, Section 7.4.2, and Section 8. After Cycle 1, nivolumab or ipilimumab may be modified as outlined in Section 7.4.1, Section 7.4.2, and Section 8.

In Arm C, sunitinib will be administered at Cycle 1. After Cycle 1, modifications in sunitinib dosing (delay, reduction/escalation, and discontinuation) may occur as outlined in Section 7.4.1, Section 7.4.2, and Section 8.

7.4.1 Dose Delay Criteria

Dose delay criteria for management of adverse events during nivolumab, ipilimumab, cabozantinib, or sunitinib treatment are outlined in this section.

Dosing of nivolumab (in Arm A) or nivolumab and ipilimumab (in Arm B) may be delayed without delay of cabozantinib dosing if toxicity is felt to be related to only nivolumab (Arm A) or nivolumab and ipilimumab (Arm B) and not related to cabozantinib. Conversely, dosing of cabozantinib may be delayed without delay of nivolumab dosing (in Arm A) or nivolumab and ipilimumab dosing (in Arm B) if toxicity is felt to be related to only cabozantinib and not related to nivolumab (Arm A) or nivolumab and ipilimumab (Arm B). However, if toxicity is considered related to all study drugs or if the investigator is unable to determine which study drug is the cause of the AE, then all study drugs in the combination should be delayed.

Participants who require dose delay should be re-evaluated weekly or more frequently if clinically indicated and resume dosing when re-treatment criteria are met (see Section 7.4.3).

7.4.1.1 Dose Delay Criteria for Nivolumab and Ipilimumab

In Arm A, nivolumab administration should be delayed, and in Arm B, both nivolumab and ipilimumab administration should be delayed, for any of the following:

- Any Grade ≥ 2 non-skin, drug-related adverse event, with the exception of fatigue
- Grade 2 drug-related creatinine, AST, ALT, and/or total bilirubin abnormalities
- Any Grade 3 skin, drug-related adverse event
- Any Grade 3 drug-related laboratory abnormality, with the following exceptions:
  - Grade 3 lymphopenia does not require dose delay
  - Grade 3 amylase or lipase abnormalities that are not associated with symptoms or clinical manifestations of pancreatitis do not require a dose delay.
  - Grade ≥ 3 AST, ALT or total bilirubin will require dose discontinuation
- Any adverse event, laboratory abnormality, or intercurrent illness which, in the judgment of the investigator, warrants delaying the dose of study medication.

During Cycles 1-4 in Arm B, both nivolumab and ipilimumab must be delayed at the same time.

Immuno-oncology agents, such as nivolumab and ipilimumab, are associated with AEs that differ in severity and duration than AEs caused by other therapeutic classes. Early recognition and management of AEs associated with immuno-oncology agents may mitigate severe toxicity. Management Algorithms have been developed to assist investigators in assessing and managing the following groups of AEs:

- Gastrointestinal
- Renal
- Pulmonary
- Hepatic
- Endocrinopathy
- Skin
- Neurological

The management algorithms recommended for use in CA2099ER are included in Appendix 5.

7.4.1.2 Dose Delay Criteria for Cabozantinib

Cabozantinib dosing should be delayed for the following:

- Urine protein/creatinine ratio (UPCR) ≥ 2.0 *or* urine dipstick protein ≥ 3+ *or* urine protein ≥. 2.0 g / 24 hours. Obtain 24 hour urine protein prior to next dosing visit.
- Grade 3 prolonged QTc interval (ie, QTcF interval > 500 msec on at least 2 out of 3 separate ECGs performed at least 3 minutes apart)
- Any Grade 2 or 3 drug-related venous thrombosis requiring anticoagulation, with the following exception:
  - Any recurrent or worsening venous thromboembolic event after restarting cabozantinib will require discontinuation
- Any other Grade 2 drug-related adverse event that persists or worsens despite supportive care management, with the following exception:

  - Any Grade 2 drug-related hemorrhage requires dose delay
  - Grade ≥ 2 arterial thromboembolic events, including cerebrovascular ischemia, cardiac ischemia/infarction, peripheral or visceral arterial ischemia, require discontinuation

- Sustained Grade 3 drug-related hypertension (systolic BP ≥ 160 mm Hg or diastolic BP ≥ 100 mm Hg). NOTE: Stopping or reducing the dose of cabozantinib is expected to cause a decrease in BP. The treating physician should monitor the participant for hypotension and adjust the number and dose of antihypertensive medications accordingly.
- Any other Grade 3 drug-related adverse event or laboratory abnormality, with the following exceptions:
  - Grade ≥ 3 drug-related hemorrhage requires discontinuation (Section 8.1.2).
  - Drug-related AST or ALT > 8 x ULN requires discontinuation (Section 8.1.2).
  - Concurrent AST or ALT > 3 x ULN and total bilirubin > 2 x ULN, consistent with potential drug-induced liver injury (see Section 9.2.7), requires discontinuation (Section 8.1.2)
  - Grade 3 drug-related amylase or lipase abnormalities that are not associated with symptoms or clinical manifestations of pancreatitis do not require dose delay. In such cases, more frequent monitoring (eg, weekly) of amylase and lipase is recommended. If amylase or lipase worsen to Grade 4 severity or the participant develops symptoms or clinical manifestations of pancreatitis, dosing should be delayed.
- Grade 4 drug-related amylase or lipase abnormalities require dose delay. Participants should be monitored for development of symptoms or clinical manifestations of pancreatitis.
- Grade 4 drug-related electrolyte abnormalities require dose delay. Electrolyte correction with supplementation/appropriate management should be promptly initiated.
- Grade 4 drug-related neutropenia, lymphopenia, leukopenia, anemia, or thrombocytopenia
- Any adverse event, laboratory abnormality, or intercurrent illness which, in the judgment of the investigator, warrants delaying cabozantinib dosing
- For participants scheduled for major surgery, including dental surgery which may impact bone healing, cabozantinib dosing should be delayed at least 28 days prior to scheduled surgery. The treating physician should use clinical judgment with regard to the risks and benefits of the planned surgical procedure if it is not possible to delay cabozantinib dosing for 28 days prior to the procedure. A delay of cabozantinib dosing of 5 to 7 days is recommended for healing for minor surgery.

As a general approach, all AEs related to cabozantinib should be managed with supportive care when possible at the earliest signs of toxicity. Calcium, magnesium, potassium, and phosphorus should be kept above the lower limits of the laboratory normal values. Please refer to the

cabozantinib IB, Appendix 10, and Appendix 11 for additional information regarding dose modifications and AE management.[22]

7.4.1.3 Dose Delay Criteria for Sunitinib

Sunitinib dose delays should be based on instructions in the approved product label and should be considered for any severe or intolerable drug-related adverse events.

Within a cycle, missed doses of sunitinib should be skipped. Participants should never be dosed during the 2-week off period of each 6-week cycle, even if treatment delays occurred earlier in the cycle and therapy is ready to be resumed. If treatment is delayed past the end of the 6-week cycle, the start of the next cycle should be delayed until treatment with sunitinib resumes.

Prior to resuming therapy after a dose delay, refer to Section 7.4.2 for dose reduction recommendations and Section 8 for discontinuation criteria.

For this protocol, sunitinib dosing should be delayed for any of the following:

- Urine protein/creatinine ratio (UPCR) ≥ 2.0 *or* urine dipstick protein ≥ 3+ *or* urine protein ≥ 2.0 g / 24 hours. Obtain 24 hour urine protein prior to next dosing visit.
- Grade 3 prolonged QTc interval (ie, QTcF interval > 500 msec on at least 2 out of 3 separate ECGs performed at least 3 minutes apart)
- Any Grade 2 or 3 drug-related venous thrombosis requiring anticoagulation, with the following exception:
    - Any recurrent or worsening venous thromboembolic event after restarting sunitinib will require discontinuation
- Any other Grade 2 drug-related adverse event that persists or worsens despite supportive care management, with the following exception:
    - Any Grade 2 drug-related hemorrhage requires dose delay
    - Grade ≥ 2 arterial thromboembolic events, including cerebrovascular ischemia, cardiac ischemia/infarction, peripheral or visceral arterial ischemia, require discontinuation
- Grade 3 drug-related hypertension (systolic BP ≥ 160 mm Hg or diastolic BP ≥ 100 mm Hg). NOTE: Stopping or reducing the dose of sunitinib is expected to cause a decrease in BP. The treating physician should monitor the participant for hypotension and adjust the number and dose of antihypertensive medications accordingly.
- Any other Grade 3 drug-related adverse event or laboratory abnormality, with the following exceptions:
    - Grade ≥ 3 drug-related hemorrhage requires discontinuation (Section 8.1.2).
    - Drug-related AST or ALT > 8 x ULN requires discontinuation (Section 8.1.2).
    - Concurrent AST or ALT > 3 x ULN and total bilirubin > 2 x ULN, consistent with potential drug-induced liver injury (see Section 9.2.7), requires discontinuation (Section 8.1.2)

- Grade 3 drug-related amylase or lipase abnormalities that are not associated with symptoms or clinical manifestations of pancreatitis do not require dose delay. In such cases, more frequent monitoring (eg, weekly) of amylase and lipase is recommended. If amylase or lipase worsen to Grade 4 severity or the participant develops symptoms or clinical manifestations of pancreatitis, dosing should be delayed.

- Grade 4 drug-related amylase or lipase abnormalities require dose delay. Participants should be monitored for development of symptoms or clinical manifestations of pancreatitis.
- Grade 4 drug-related electrolyte abnormalities require dose delay. Electrolyte correction with supplementation/appropriate management should be promptly initiated.
- Grade 4 drug-related neutropenia, lymphopenia, leukopenia, anemia, or thrombocytopenia
- Any adverse event, laboratory abnormality, or intercurrent illness which, in the judgment of the investigator, warrants delaying cabozantinib dosing
- For participants scheduled for major surgery, including dental surgery which may impact bone healing, sunitinib dosing should be delayed at least 28 days prior to scheduled surgery. The treating physician should use clinical judgment with regard to the risks and benefits of the planned surgical procedure if it is not possible to delay sunitinib dosing for 28 days prior to the procedure. A delay of sunitinib dosing of 5 to 7 days is recommended for healing for minor surgery.

As a general approach, all AEs related to sunitinib should be managed with supportive care when possible at the earliest signs of toxicity. Calcium, magnesium, potassium, and phosphorus should be kept above the lower limits of the laboratory normal values. Please refer to the sunitinib approved product label and Appendix 10 for additional information regarding AE management.

7.4.2 Dose Reductions and Escalations

Dose reductions are permitted for cabozantinib and sunitinib but not for nivolumab and ipilimumab (see Table 7.4.2-1).

Table 7.4.2-1: Dose Level Modifications Table

Dose Level	**Cabozantinib (tablet dose expression)**	**Sunitinib (capsule dose expression)**	**Nivolumab (IV)**	**Ipilimumab (IV)**	**Nivolumab (IV)**
0 (starting dose)	40 mg daily	50 mg daily	3mg/kg	1 mg/kg	240 mg
-1	20 mg daily	37.5 mg daily	-	-	-
-2	20 mg every other day	25 mg daily	-	-	-

7.4.2.1 Dose Reduction and Escalation for Nivolumab and Ipilimumab

No dose reductions or dose escalations of nivolumab or ipilimumab are allowed.

7.4.2.2 Dose Reduction and Escalation for Cabozantinib

Dose reductions and dose escalations for adverse event management are allowed for cabozantinib (Table 7.4.2-1). Cabozantinib doses will not be re-escalated once reduced, unless a concomitant strong CYP3A4 inducer is started (see below). The only exception is participants who continue on cabozantinib alone after discontinuation of nivolumab alone (Arm A) or discontinuation of both nivolumab and ipilimumab (Arm B) who may re-escalate one dose level if the prior dose reduction was due to a toxicity felt by the investigator to have been mainly related to nivolumab and/or ipilimumab.

After toxicity requiring a dose delay has improved and meets the criteria to resume dosing (Section 7.4.3.2), participants who were receiving cabozantinib 40 mg daily prior to the delay will resume cabozantinib at 20 mg daily. Participants who were receiving cabozantinib 20 mg daily prior to the delay and require another dose delay will resume cabozantinib at 20 mg every other day. If more than 2 dose reductions are necessary (ie, reduction to less than 20 mg every other day), cabozantinib must be permanently discontinued (Section 8.1.2).

Participants who required a dose delay due to Grade 3 hypertension, which improved with antihypertensive medications, or any Grade 2 or 3 drug-related adverse event or asymptomatic laboratory abnormality that improved to Grade ≤ 1 within 7 days with supportive medical care may resume cabozantinib at the same dose or a reduced dose, at the discretion of the investigator.

Participants with Grade 3 drug-related lipase or amylase abnormalities that are not associated with symptoms or clinical manifestations of pancreatitis may reduce cabozantinib by one dose level, without delaying dosing, at the discretion of the investigator.

In Participants Who Start Taking a Concomitant Strong CYP3A4 Inhibitor: Reduce the daily cabozantinib dose by 20 mg (for example, from 40 mg to 20 mg daily). Resume the dose that was used prior to initiating the CYP3A4 inhibitor 2 to 3 days after discontinuation of the strong inhibitor (see Prescribing Information for cabozantinib).

In Participants Who Start Taking a Concomitant Strong CYP3A4 Inducer: Increase the daily cabozantinib dose by 20 mg (eg, from 40 mg to 60 mg daily) as tolerated. Resume the dose that was used prior to initiating the CYP3A4 inducer 2 to 3 days after discontinuation of the strong inducer. In this study, the daily dose of cabozantinib should not exceed 60 mg (see Prescribing Information for cabozantinib).

7.4.2.3 Dose Reduction and Escalation for Sunitinib

Sunitinib Dose Reductions are permitted as per the approved product label for safety reasons or when a concomitant strong CYP3A4 inhibitor is needed (Appendix 9). Selection of an alternative concomitant medication with minimal or no enzyme inhibition potential is recommended whenever possible.

After toxicity requiring a dose delay has improved and meets the criteria to resume dosing (Section 7.4.3.3), participants should resume sunitinib at one dose level reduction. Dose reductions should occur in 12.5 mg decrements. No more than 2 dose reductions are allowed. If

more than 2 dose reductions are necessary (ie, reduction to less than 25 mg daily), the participant must be permanently discontinued (Section 8.1.3).

Participants who required a dose delay due to Grade 3 hypertension, which improved with antihypertensive medications, or any Grade 2 or 3 drug-related adverse event or asymptomatic laboratory abnormality that improved to Grade ≤ 1 within 7 days with supportive medical care may resume sunitinib at the same dose or a reduced dose, at the discretion of the investigator.

At the time a dose reduction is considered, also refer to Section 7.4.2.3 for dose delay recommendations and Section 8.1.3 for discontinuation criteria.

Sunitinib Dose Escalations are permitted as per the approved product label when a concomitant CYP3A4 inducer is needed (Appendix 9). Selection of an alternative concomitant medication with minimal or no enzyme induction potential is recommended whenever possible.

7.4.3 Criteria to Resume Treatment

7.4.3.1 Criteria to Resume Nivolumab and Ipilimumab Treatment

Delayed doses of nivolumab and/or ipilimumab should be administered as soon as the participant meets criteria to resume treatment. If a dose has been delayed, the participant should not wait until the next scheduled dosing date.

Participants may resume treatment with study drug when the drug-related AE(s) resolve to Grade ≤ 1 or baseline value, with the following exceptions:

- Participants may resume treatment in the presence of Grade 2 fatigue.
- Participants who have not experienced a Grade 3 drug-related skin AE may resume treatment in the presence of Grade 2 skin toxicity.
- For participants with Grade 2 AST, ALT, and/or total bilirubin abnormalities, dosing may resume when laboratory values return to baseline and management with corticosteroids, if needed, has been completed.
- Participants with combined Grade 2 AST/ALT AND total bilirubin values meeting discontinuation parameters (Section 8.1.1) should have treatment permanently discontinued.
- Drug-related pulmonary toxicity, diarrhea, or colitis, must have resolved to baseline before treatment is resumed. Participants with persistent Grade 1 pneumonitis after completion of a steroid taper over at least 1 month may be eligible for retreatment if discussed with and approved by the BMS Medical Monitor or designee.
- Drug-related endocrinopathies adequately controlled with only physiologic hormone replacement may resume treatment.
- If treatment is delayed > 6 weeks, the participant must be permanently discontinued from study therapy, except as specified in Section 8.1.1.

For Arm B participants who delay nivolumab and ipilimumab dosing after Cycle 1 or 2, both nivolumab and ipilimumab must be resumed on the same day when the criteria to resume treatment are met.

For Arm B participants who delay nivolumab and ipilimumab dosing after Cycle 3 due to any drug-related AE meeting dose delay criteria that does not resolve within 14 days or requires

treatment with systemic corticosteroids, it is acceptable to omit Cycle 4 if the investigator feels that ipilimumab was the main cause of the toxicity requiring dose delay. In this situation, when the participant meets criteria to resume nivolumab, the participant may proceed to Cycle 5 and begin nivolumab monotherapy maintenance.

7.4.3.2 Criteria to Resume Cabozantinib Treatment

Participants may resume treatment with cabozantinib when the drug-related AE(s) resolve to Grade ≤ 1 or baseline value, with the following exceptions:

- Participants may resume dosing in the presence of Grade 2 fatigue
- Participants who delayed dosing due to prolonged QTcF may resume dosing at one reduced dose level once QTcF returns to ≤ 500 msec
- Participants who delayed dosing due to UPCR ≥ 2.0 or urine dipstick protein ≥ 3+ may resume dosing at one dose level reduction when 24 hour urine protein < 2.0 g
- Participants who delayed dosing due to Grade 3 hypertension may resume dosing at the same dose or at one dose level reduction, at the discretion of the investigator, when hypertension has improved to Grade ≤ 2
- Participants who delayed dosing due to Grade 4 lipase or amylase abnormalities may resume dosing upon resolution to Grade ≤ 2
- Participants who delayed dosing due to major surgery should not resume cabozantinib until complete wound healing has taken place. Following cabozantinib resumption, participants should be monitored for wound dehiscence, wound infections, and other signs of impaired wound healing.
- If treatment is delayed > 6 weeks for any reason, the participant must be permanently discontinued from study therapy, except in cases where permission to resume treatment is granted by the BMS Medical Monitor or designee.

7.4.3.3 Criteria to Resume Sunitinib Treatment

Within a cycle, missed doses of sunitinib should be skipped and not replaced. Participants should never be dosed during the 2-week off period of each 6-week cycle, even if treatment delays occurred earlier in the cycle and therapy is ready to be resumed. If treatment is delayed past the end of the 6-week cycle, the start of the next cycle should be delayed until treatment with sunitinib resumes.

Participants may resume treatment with sunitinib when the drug-related AE(s) resolve to Grade ≤ 1 or baseline value, with the following exceptions:

- Participants may resume dosing in the presence of Grade 2 fatigue
- Participants who delayed dosing due to prolonged QTcF may resume dosing at one reduced dose level once QTcF returns to ≤ 500 msec
- Participants who delayed dosing due to UPCR ≥ 2.0 or urine dipstick protein ≥ 3+ may resume dosing at one dose level reduction when 24 hour urine protein < 2.0 g
- Participants who delayed dosing due to Grade 3 hypertension may resume dosing at the same dose or at one dose level reduction, at the discretion of the investigator, when hypertension has improved to Grade ≤ 2

- Participants who delayed dosing due to Grade 4 lipase or amylase abnormalities may resume dosing upon resolution to Grade ≤ 2
- If treatment is delayed > 6 weeks for any reason, the participant must be permanently discontinued from study therapy, except in cases where permission to resume treatment is granted by the BMS Medical Monitor or designee.

7.4.4 Treatment of Nivolumab- or Ipilimumab-Related Infusion Reactions

Since nivolumab and ipilimumab contain only human immunoglobulin protein sequences, they are unlikely to be immunogenic and induce infusion or hypersensitivity reactions. However, if such a reaction were to occur, it might manifest with fever, chills, rigors, headache, rash, pruritus, arthralgias, hypotension, hypertension, bronchospasm, or other allergic-like reactions. Regardless of whether or not the event is attributed to the study drugs, all Grade 3 or 4 infusion reactions should be reported within 24 hours to the study BMS Medical Monitor or designee and reported as an SAE if it meets the criteria, given the blinded nature of the study. Infusion reactions should be graded according to NCI CTCAE (Version 4) guidelines.

Treatment recommendations are provided below and may be modified based on local treatment standards and guidelines, as appropriate:

For Grade 1 symptoms: (mild reaction; infusion interruption not indicated; intervention not indicated):

- Remain at bedside and monitor participant until recovery from symptoms. The following prophylactic premedications are recommended for future infusions: diphenhydramine 50 mg (or equivalent) and/or acetaminophen/paracetamol 325 to 1000 mg at least 30 minutes before additional study drug administrations.

For Grade 2 symptoms: (moderate reaction required therapy or infusion interruption but responds promptly to symptomatic treatment (eg, antihistamines, non-steroidal anti-inflammatory drugs, narcotics, corticosteroids, bronchodilators, IV fluids); prophylactic medications indicated for ≤ 24 hours):

- Stop the study drug infusion, begin an IV infusion of normal saline, and treat the participant with diphenhydramine 50 mg IV (or equivalent) and/or acetaminophen/paracetamol 325 to 1000 mg; remain at bedside and monitor participant until resolution of symptoms. Corticosteroid and/or bronchodilator therapy may also be administered as appropriate. If the infusion is interrupted, then restart the infusion at 50% of the original infusion rate when symptoms resolve; if no further complications ensue after 30 minutes, the rate may be increased to 100% of the original infusion rate. Monitor participant closely. If symptoms recur, then no further study dug will be administered at that visit.
- For future infusions, the following prophylactic premedications are recommended: diphenhydramine 50 mg (or equivalent) and/or acetaminophen/paracetamol 325 to 1000 mg should be administered at least 30 minutes before study drug infusions. If necessary, corticosteroids (up to 25 mg of hydrocortisone or equivalent) may be used.

For Grade 3 or 4 symptoms: (severe reaction, Grade 3: prolonged [ie, not rapidly responsive to symptomatic medication and/or brief interruption of infusion]; recurrence of symptoms following initial improvement; hospitalization indicated for other clinical sequelae (eg, renal

impairment, pulmonary infiltrates). Grade 4: Life-threatening; pressor or ventilatory support indicated):

- Immediately discontinue infusion of study drug. Begin an IV infusion of normal saline and treat the participant as follows: Recommend bronchodilators, epinephrine 0.2 to 1 mg of a 1:1000 solution for subcutaneous administration or 0.1 to 0.25 mg of a 1:10,000 solution injected slowly for IV administration, and/or diphenhydramine 50 mg IV with methylprednisolone 100 mg IV (or equivalent), as needed. Participant should be monitored until the Investigator is comfortable that the symptoms will not recur. Study drug will be permanently discontinued. Investigators should follow their institutional guidelines for the treatment of anaphylaxis. Remain at bedside and monitor participant until recovery of the symptoms.

In case of late-occurring hypersensitivity symptoms (eg, appearance of a localized or generalized pruritus within 1 week after treatment), symptomatic treatment may be given be given (eg, oral antihistamine or corticosteroids).

7.5 Preparation/Handling/Storage/Accountability

The investigational product should be stored in a secure area according to local regulations. It is the responsibility of the investigator to ensure that investigational product is only dispensed to study Participants. The investigational product must be dispensed only from official study sites by authorized personnel according to local regulations.

The product storage manager should ensure that the study treatment is stored in accordance with the environmental conditions (temperature, light, and humidity) as determined by BMS. If concerns regarding the quality or appearance of the study treatment arise, the study treatment should not be dispensed and contact BMS immediately.

Study treatment not supplied by BMS will be stored in accordance with the package insert.

Investigational product documentation (whether supplied by BMS or not) must be maintained that includes all processes required to ensure drug is accurately administered. This includes documentation of drug storage, administration and, as applicable, storage temperatures, reconstitution, and use of required processes (eg, required diluents, administration sets).

- Further guidance and information for final disposition of unused study treatment are provided in Appendix 2.

7.5.1 Retained Samples for Bioavailability / Bioequivalence

Not applicable.

7.6 Treatment Compliance

Study treatment compliance will be periodically monitored by drug accountability (including review of dosing diary cards, as applicable). Drug accountability should be reviewed by the site study staff at each visit to confirm treatment compliance. Sites should discuss discrepancies with the participant at each on-treatment study visit.

7.7 Concomitant Therapy

7.7.1 Prohibited and/or Restricted Treatments

The following medications are prohibited during the study (unless utilized to treat a drug related adverse event):

- Immunosuppressive agents
- Immunosuppressive doses of systemic corticosteroids (except as stated in Section 7.7.3)
- Any concurrent anti-neoplastic therapy (ie, chemotherapy, hormonal therapy, immunotherapy, extensive, non-palliative radiation therapy, or standard or investigational agents)

Supportive care for disease-related symptoms may be offered to all participants on the trial.

7.7.2 Other Restrictions and Precautions

Participants with a condition requiring systemic treatment with either corticosteroids (> 10 mg daily prednisone equivalent) or other immunosuppressive medications within 14 days of randomization are excluded. Inhaled or topical steroids, and adrenal replacement steroid doses > 10 mg daily prednisone equivalent, are permitted in the absence of active autoimmune disease.

7.7.2.1 Restrictions and Precautions for Cabozantinib

Cabozantinib is a substrate of CYP3A4. In Arm A and Arm B participants, co-administration of cabozantinib with medications that are strong inhibitors/inducers of CYP3A4 should be avoided.

Cabozantinib is highly protein bound (99.9%) to human plasma proteins. Concomitant medications that are highly protein bound (eg, diazepam, furosemide, dicloxacillin, propanolol) should be used with caution. Because warfarin is a highly protein bound drug with a low therapeutic index, administration of warfarin should be avoided in participants receiving cabozantinib in Arms A and B.

Cabozantinib has been associated with a mild prolongation of the QTc interval. Caution should be used when treating participants on cabozantinib in Arms A and B with other drugs associated with QTc prolongation. Additional QTc monitoring is suggested for participants who are treated concomitantly with QTc prolonging drugs.

Low-dose aspirin for cardioprotection (as per local applicable guidelines) and low-dose low molecular weight heparins (LMWH) are permitted. Anticoagulation with therapeutic doses of LMWH is allowed in participants without known brain metastases who are on a stable dose of LMWH for at least 6 weeks before first dose of study treatment, and who have had no clinically significant hemorrhagic complications from the anticoagulation regimen or the tumor.

While on study, invasive dental procedures (eg, extractions, dental implants) should be avoided and conservative dental therapy, such as endodontic therapy (root canal treatment) would be preferred. After a surgical dental procedure, allow complete wound healing before resuming with cabozantinib treatment.

7.7.2.2 Restrictions and Precautions for Sunitinib

Sunitinib is metabolized by CYP3A4. Arm C participants who initiate CYP3A4 inhibitors/inducers after randomization should follow sunitinib dose modification instructions.

Sunitinib has been shown to prolong the QTc interval. Concomitant treatment with dysrhythmic drugs (ie, terfenadine, quinidine, procainamide, sotalol, probucol, bepridil, haloperidol, risperidone, and indapamide) is not recommended.

Medications taken within 4 weeks prior to study drug administration must be recorded on the CRF.

7.7.2.3 Imaging Restriction and Precautions

It is the local imaging facility's responsibility to determine, based on participant attributes (eg, allergy history, diabetic history and renal status), the appropriate imaging modality and contrast regimen for each participant. Imaging contraindications and contrast risks should be considered in this assessment. Participants with renal insufficiency should be assessed as to whether or not they should receive contrast and if so, what type and dose of contrast is appropriate. Specific to MRI, participants with severe renal insufficiency (ie, estimated glomerular filtration rate [eGFR] < 30 mL/min/1.73 m^2) are at increased risk of nephrogenic systemic fibrosis. MRI contrast should not be given to this participant population. In addition, participants are contraindicated from MRI if they have tattoos, metallic implants, pacemakers, etc.

The ultimate decision to perform MRI in an individual participant in this study rests with the site radiologist, the investigator and the standard set by the local Ethics Committee.

7.7.3 Permitted Therapy

Participants are permitted the use of topical, ocular, intra-articular, intranasal, and inhalational corticosteroids (with minimal systemic absorption). Adrenal replacement steroid doses > 10 mg daily prednisone are permitted. A brief (less than 3 weeks) course of corticosteroids for prophylaxis (eg, contrast dye allergy) or for treatment of non-autoimmune conditions (eg, delayed-type hypersensitivity reaction caused by a contact allergen) is permitted.

Concomitant medications are recorded at baseline and throughout the treatment phase of the study in the appropriate section of the CRF. All medications (prescriptions or over the counter medications) continued at the start of the study or started during the study and different from the study drug must be documented in the concomitant therapy section of the CRF.

7.8 Treatment After the End of the Study

At the conclusion of the study, participants who continue to demonstrate clinical benefit will be eligible to receive BMS supplied study treatment for the maximum treatment duration specified in protocol Section 7.1. Study treatment will be provided via an extension of the study, a rollover study requiring approval by responsible health authority and ethics committee or through another mechanism at the discretion of BMS.

BMS reserves the right to terminate access to BMS supplied study treatment if any of the following occur: a) the study is terminated due to safety concerns; b) the development of the nivolumab, ipilimumab, cabozantinib, or sunitinib is terminated for other reasons, including but not limited to lack of efficacy and/or not meeting the study objectives; c) the participant can obtain medication from a government sponsored or private health program. In all cases BMS will follow local regulations.

8. DISCONTINUATION CRITERIA

8.1 Discontinuation from Study Treatment

Participants MUST discontinue investigational product (and non-investigational product at the discretion of the investigator) for any of the following reasons:

- Participant's request to stop study treatment. Participants who request to discontinue study treatment will remain in the study and must continue to be followed for protocol specified follow-up procedures. The only exception to this is when a participant specifically withdraws consent for any further contact with him/her or persons previously authorized by participant to provide this information
- Any clinical adverse event (AE), laboratory abnormality or intercurrent illness which, in the opinion of the investigator, indicates that continued participation in the study is not in the best interest of the participant
- Termination of the study by Bristol-Myers Squibb (BMS)
- Loss of ability to freely provide consent through imprisonment or involuntarily incarceration for treatment of either a psychiatric or physical (eg, infectious disease) illness
- Criteria listed in 8.1.1
- Disease progression of RCC or occurrence of a secondary malignancy which requires systemic therapy for treatment

Refer to the Schedule of Activities for data to be collected at the time of treatment discontinuation and follow-up and for any further evaluations that can be completed.

In the case of pregnancy, the investigator must immediately notify the BMS Medical Monitor/designee of this event. In most cases, the study treatment will be permanently discontinued in an appropriate manner (eg, dose tapering if necessary for participant safety). Please call the BMS Medical Monitor within 24 hours of awareness of the pregnancy. If the investigator determines a possible favorable benefit/risk ratio that warrants continuation of study treatment, a discussion between the investigator and the BMS Medical Monitor/designee must occur.

All participants who discontinue study treatment should comply with protocol specified follow-up procedures as outlined in Section 2. The only exception to this requirement is when a participant withdraws consent for all study procedures including post-treatment study follow-up or loses the ability to consent freely (ie, is imprisoned or involuntarily incarcerated for the treatment of either a psychiatric or physical illness).

If study treatment is discontinued prior to the participant's completion of the study, the reason for the discontinuation must be documented in the participant's medical records and entered on the appropriate case report form (CRF) page.

8.1.1 Nivolumab and Ipilimumab Dose Discontinuation (Arms A and B)

The assessment for discontinuation of nivolumab, ipilimumab, and cabozantinib should be made separately for each study drug. Although there is overlap among the discontinuation criteria, if discontinuation criteria are met for one study drug but not the other(s), it may be acceptable to continue treatment with the study drug(s) that are not felt to be related the toxicity, as specified below. If the investigator considers the toxicity to be related to all study drugs or is unable to determine which of the study drug(s) in Arm A or Arm B are the cause of toxicity, then all study drugs in the treatment regimen should be discontinued, and the recommendations for management of toxicity related to all study drugs should be promptly initiated.

Nivolumab and/or ipilimumab treatment should be permanently discontinued for any of the following:

- Any Grade 2 drug-related uveitis, eye pain or blurred vision that does not respond to topical therapy and does not improve to Grade 1 severity within the re-treatment period OR requires systemic treatment
- Any Grade 3 non-skin, drug-related adverse event lasting > 7 days, or recurs with the following exceptions for laboratory abnormalities, diarrhea, colitis, neurologic toxicity, drug-related uveitis, pneumonitis, bronchospasm, hypersensitivity reactions, infusion reactions, and endocrinopathies:
    - Grade 3 drug-related diarrhea, colitis, neurologic toxicity, uveitis, pneumonitis, bronchospasm, hypersensitivity reaction, or infusion reaction of any duration requires discontinuation. NOTE: The diagnosis of colitis should be supported by findings on colonoscopy whenever possible.
    - Grade 3 drug-related endocrinopathies, adequately controlled with only physiologic hormone replacement do not require discontinuation. Adrenal insufficiency requires discontinuation regardless of control with hormone replacement.
    - Grade 3 drug-related laboratory abnormalities do not require treatment discontinuation except:
        - Grade 3 drug-related thrombocytopenia > 7 days or associated with bleeding requires discontinuation
        - Any drug-related liver function test (LFT) abnormality that meets the following criteria require discontinuation:
            - Grade ≥ 3 drug-related AST, ALT or Total Bilirubin requires discontinuation*
            - Concurrent AST or ALT > 3 x ULN and total bilirubin > 2x ULN

* In most cases of Grade 3 AST or ALT elevation, study drug(s) will be permanently discontinued. If the investigator determines a possible favorable benefit/risk ratio that warrants continuation of study drug(s), a discussion between the investigator and the BMS Medical Monitor/designee must occur.

- Any Grade 4 drug-related adverse event or laboratory abnormality (including but not limited to creatinine, AST, ALT, or Total Bilirubin), except for the following events which do not require discontinuation:
  - Grade 4 neutropenia ≤ 7 days
  - Grade 4 lymphopenia or leukopenia or asymptomatic amylase or lipase
  - Isolated Grade 4 electrolyte imbalances/abnormalities that are not associated with clinical sequelae and are corrected with supplementation/appropriate management within 72 hours of their onset
  - Grade 4 drug-related endocrinopathy adverse events, such as, hyper- or hypothyroidism, or glucose intolerance, which resolve or are adequately controlled with physiologic hormone replacement (corticosteroids, thyroid hormones) or glucose-controlling agents, respectively, may not require discontinuation after discussion with and approval from the BMS Medical Monitor or designee.
- Any event that leads to delay in dosing lasting > 6 weeks from the previous dose requires discontinuation, with the following exceptions:
  - Dosing delays to allow for prolonged steroid tapers to manage drug-related adverse events are allowed.
  - Dosing delays lasting > 6 weeks from the previous dose that occur for non-drug-related reasons may be allowed if approved by the BMS Medical Monitor or designee.

Prior to re-initiating treatment in a participant with a dosing delay lasting > 6 weeks, the BMS Medical Monitor or designee must be consulted. Tumor assessments should continue as per protocol even if dosing is delayed. Periodic study visits to assess safety and laboratory studies should also continue every 6 weeks or more frequently if clinically indicated during such dosing delays.

- Any adverse event, laboratory abnormality, or intercurrent illness which, in the judgment of the Investigator, presents a substantial clinical risk to the participant with continued nivolumab dosing.

Participants in Arm B who meet any of the discontinuation criteria above prior to Cycle 3 must discontinue both nivolumab and ipilimumab and may not receive nivolumab monotherapy maintenance.

Participants in Arm B who meet any of the discontinuation criteria above after Cycle 3 or Cycle 4 may be able to proceed to Cycle 5 (skipping Cycle 4 if needed) to begin nivolumab

monotherapy maintenance if the toxicity is felt to be related mainly to ipilimumab, only after discussion with and approval by the BMS Medical Monitor or designee.

8.1.2 Cabozantinib Dose Discontinuation

Permanently discontinue cabozantinib for participants with any of the following:

- Any requirement for more than 2 cabozantinib dose reductions (ie, reduction to less than 20 mg every other day)
- Any Grade ≥ 2 drug-related arterial thromboembolic events, including cerebrovascular ischemia, cardiac ischemia/infarction, peripheral or visceral arterial ischemia
- Any Grade ≥ 3 drug-related hemorrhage
- Grade 4 hypertension or persistent Grade 3 hypertension despite optimal medical management and cabozantinib dose reduction
- Drug-related reversible posterior leukoencephalopathy syndrome
- Development of drug-related fistula or GI perforation
- Drug-related nephrotic syndrome
- Any drug-related liver function test (LFT) abnormality that meets the following criteria requires discontinuation:
    - AST or ALT > 8 x ULN.
    - Concurrent AST or ALT > 3 x ULN and total bilirubin >2 x ULN, consistent with potential drug-induced liver injury (see Section 9.2.7).
- Any Grade 4 drug-related adverse event or laboratory abnormality, with the following exceptions:
    - Grade 4 neutropenia ≤ 7 days
    - Grade 4 lymphopenia or leukopenia
    - Grade 4 amylase or lipase abnormalities that are not associated with symptoms or clinical manifestations of pancreatitis
    - Isolated Grade 4 electrolyte abnormalities that are not associated with clinical sequelae and are corrected with supplementation/appropriate management within 72 hours of their onset
- Any dosing interruption lasting > 6 weeks unless the BMS Medical Monitor or designee is consulted and agrees with the rationale for resuming therapy after a delay > 6 weeks. Note that tumor assessments should continue as per protocol even if dosing is interrupted.
- Any adverse event, laboratory abnormality, or intercurrent illness which, in the judgment of the Investigator, presents a substantial clinical risk to the participant with continued sunitinib dosing.

8.1.3 Sunitinib Dose Discontinuation

Treatment with sunitinib should be permanently discontinued for any of the following:

- Any requirement for more than 2 sunitinib dose reductions (ie, reduction to less than 25 mg daily)
- Any Grade drug-related arterial thrombosis.
- Grade 4 drug-related hemorrhage or recurrent Grade 3 drug-related hemorrhage after dose reduction.
- Grade 4 drug-related symptomatic venous thrombosis.
- Grade 4 drug-related cardiac toxicity.
- Grade 4 hypertension or persistent Grade 3 hypertension despite optimal medical management and sunitinib dose reduction
- Any drug-related liver function test (LFT) abnormality that meets the following criteria requires discontinuation:
    - AST or ALT > 8 x ULN.
    - Concurrent AST or ALT > 3 x ULN and total bilirubin >2 x ULN, consistent with potential drug-induced liver injury (see Section 9.2.7).
- Any other Grade 4 drug-related adverse event or laboratory abnormality, with the following exceptions:
    - Grade 4 neutropenia ≤ 7 days
    - Grade 4 lymphopenia or leukopenia
    - Grade 4 amylase or lipase abnormalities that are not associated with symptoms or clinical manifestations of pancreatitis
    - Isolated Grade 4 electrolyte abnormalities that are not associated with clinical sequelae and are corrected with supplementation/appropriate management within 72 hours of their onset
- Any dosing interruption lasting > 6 weeks unless the BMS Medical Monitor or designee is consulted and agrees with the rationale for resuming therapy after a delay > 6 weeks. Note that tumor assessments should continue as per protocol even if dosing is interrupted.
- Any adverse event, laboratory abnormality, or intercurrent illness which, in the judgment of the Investigator, presents a substantial clinical risk to the participant with continued sunitinib dosing.

8.1.4 Treatment Beyond Disease Progression

Accumulating evidence indicates a minority of participants treated with immunotherapy or anti-angiogenic therapy may derive clinical benefit despite initial evidence of PD.[20,50,51]

Participants, regardless of study arm, will be permitted to continue treatment beyond initial RECIST 1.1 defined PD, assessed by the investigator, as long as they meet the following criteria:

- Investigator-assessed clinical benefit
- Tolerance of study drug

- Stable performance status
- Treatment beyond progression will not delay an imminent intervention to prevent serious complications of disease progression (eg, CNS metastases)
- Participant provides written informed consent prior to receiving additional treatment with the study drug regimen. All other elements of the main consent including description of reasonably foreseeable risks or discomforts, or other alternative treatment options will still apply.

A radiographic assessment/ scan should be performed within 6 weeks of initial investigator-assessed progression to determine whether there has been a decrease in the tumor size or continued PD. The assessment of clinical benefit should be balanced by clinical judgment as to whether the participant is clinically deteriorating and unlikely to receive any benefit from continued treatment with nivolumab.

If the investigator feels that the participant continues to achieve clinical benefit by continuing treatment, the participant should remain on the trial and continue to receive monitoring according to the Schedule of Activities Schedule in Section 2.

For the participants who continue study therapy beyond progression, further progression is defined as an additional 10% increase in tumor burden with a minimum 5 mm absolute increase from time of initial PD. This includes an increase in the sum of diameters of all target lesions and/ or the diameters of new measurable lesions compared to the time of initial PD. It is recommended that study treatment should be discontinued permanently upon documentation of further progression.

New lesions are considered measureable at the time of initial progression if the longest diameter is at least 10 mm (except for pathological lymph nodes which must have a short axis of at least15 mm). Any new lesion considered non-measureable at the time of initial progression may become measureable and therefore included in the tumor burden if the longest diameter increases to at least 10 mm (except for pathological lymph nodes which must have a short axis of at least 15 mm). In situations where the relative increase in total tumor burden by 10% is solely due to inclusion of new lesions which become measurable, these new lesions must demonstrate an absolute increase of at least 5 mm.

8.1.5 Post-Study Treatment Follow-up

In this study, overall survival is a key endpoint of the study. Post-study treatment follow-up is of critical importance and is essential to preserving participant safety and the integrity of the study. Participants who discontinue study treatment must continue to be followed for collection of outcome and/or survival follow-up data as required and in line with Section 5 until death or the conclusion of the study.

BMS may request that survival data be collected on all treated/randomized participants outside of the protocol defined window (See Section 2). At the time of this request, each participant will be contacted to determine their survival status unless the participant has withdrawn consent for all contacts or is lost to follow-up.

8.2 Discontinuation from the Study

Participants who request to discontinue study treatment will remain in the study and must continue to be followed for protocol specified follow-up procedures. The only exception to this is when a participant specifically withdraws consent for any further contact with him/her or persons previously authorized by participant to provide this information.

- Participants should notify the investigator of the decision to withdraw consent from future follow-up **in writing**, whenever possible.
- The withdrawal of consent should be explained in detail in the medical records by the investigator, as to whether the withdrawal is from further treatment with study treatment only or also from study procedures and/or post treatment study follow-up, and entered on the appropriate CRF page.
- In the event that vital status (whether the participant is alive or dead) is being measured, publicly available information should be used to determine vital status only as appropriately directed in accordance with local law.
- If the participant withdraws consent for disclosure of future information, the sponsor may retain and continue to use any data collected before such a withdrawal of consent.

8.3 Lost to Follow-Up

- All reasonable efforts must be made to locate participants to determine and report their ongoing status. This includes follow-up with persons authorized by the participant.
- Lost to follow-up is defined by the inability to reach the participant after a minimum of **three** documented phone calls, faxes, or emails as well as lack of response by participant to one registered mail letter. All attempts should be documented in the participant's medical records.
- If it is determined that the participant has died, the site will use permissible local methods to obtain date and cause of death.
- If investigator's use of third-party representative to assist in the follow-up portion of the study has been included in the participant's informed consent, then the investigator may use a Sponsor retained third-party representative to assist site staff with obtaining participant's contact information or other public vital status data necessary to complete the follow-up portion of the study.
- The site staff and representative will consult publicly available sources, such as public health registries and databases, in order to obtain updated contact information.
- If after all attempts, the participant remains lost to follow-up, then the last known alive date as determined by the investigator should be reported and documented in the participant's medical records.

9. STUDY ASSESSMENTS AND PROCEDURES

- Study procedures and timing are summarized in the Schedule of Activities.
- Protocol waivers or exemptions are not allowed.
- All immediate safety concerns must be discussed with the Sponsor immediately upon occurrence or awareness to determine if the participant should continue or discontinue treatment.

- Adherence to the study design requirements, including those specified in the Schedule of Activities, is essential and required for study conduct.
- All screening evaluations must be completed and reviewed to confirm that potential participants meet all eligibility criteria before randomization. The investigator will maintain a screening log to record details of all participants screened and to confirm eligibility or record reasons for screening failure, as applicable.
- Procedures conducted as part of the participant's routine clinical management (eg, blood count) and obtained before signing of informed consent may be utilized for screening or baseline purposes provided the procedure meets the protocol-defined criteria and has been performed within the timeframe defined in the Schedule of Activities.

Additional measures, including non-study required laboratory tests, should be performed as clinically indicated or to comply with local regulations. Laboratory toxicities (eg, suspected drug induced liver enzyme evaluations) will be monitored during the follow-up phase via on site/local labs until all study drug related toxicities resolve, return to baseline, or are deemed irreversible.

If a participant shows pulmonary-related signs (hypoxia, fever) or symptoms (eg, dyspnea, cough, fever) consistent with possible pulmonary adverse events, the participant should be immediately evaluated to rule out pulmonary toxicity, according to the suspected pulmonary toxicity management algorithm in the BMS-936558 (nivolumab) Investigator Brochure.

Some of the assessments referred to in this section may not be captured as data in the eCRF. They are intended to be used as safety monitoring by the treating physician. Additional testing or assessments may be performed as clinically necessary or where required by institutional or local regulations.

9.1 Efficacy Assessments

Study evaluations will take place in accordance with the Schedule of Activities in Section 2.

Images will be submitted to an imaging core lab. Sites should be trained prior to scanning the first study participant. Image acquisition guidelines and submission process will be outlined in the CA2099ER Imaging Manual to be provided by the core lab.

9.1.1 Methods of Measurements

The following imaging assessments should be performed at pre-specified intervals: CT of the chest, CT or MRI of the abdomen, pelvis, and other known sites of disease.

- CT scans should be acquired with slice thickness of 5 mm or less with no intervening gap (continuous)
- Should a participant have a contraindication for CT IV contrast, a non-contrast CT of the chest and a contrast enhanced MRI of the abdomen and pelvis and other sites of disease may be obtained. MRIs should be acquired with slice thickness of 5 mm or less with no gap (continuous).
- Every attempt should be made to image each participant using an identical acquisition protocol on the same scanner for all imaging time points
- PET alone will not be considered for the disease assessment. Complementary CT and/or MRI or biopsy must be performed in such cases.

Note: Use of CT component of a PET/CT scanner: Combined modality scanning, such as with FDG-PET/CT, is increasingly used in clinical care and is a modality/technology that is in rapid evolution; therefore, the recommendations outlined here may change rather quickly with time. At present, low-dose or attenuation correction CT portions of a combined FDG-PET/CT are of limited use in anatomically-based efficacy assessments, and it is, therefore, suggested that they should not be substituted for dedicated diagnostic contrast enhanced CT scans for anatomically-based RECIST (Appendix 8) measurements. However, if a site can document that the CT performed as part of a FDG-PET/CT is of identical diagnostic quality to a diagnostic CT (with IV and oral contrast), then the CT portion of the FDG-PET/CT can be used for RECIST 1.1 measurements. Note, however, that the FDG-PET portion of the CT introduces additional data which may bias an investigator if it is not routinely or serially performed.

9.1.2 Imaging and Clinical Assessment

Images will be submitted to an imaging core lab. Sites should be trained prior to scanning the first study participant. Image acquisition guidelines and submission process will be outlined in the CA2099ER Imaging Manual to be provided by the core lab.

Baseline imaging, including CT/MRI of the chest, abdomen, pelvis, and all known sites of disease performed within 28 days prior to randomization, should be submitted to the imaging core lab. Baseline brain MRI (preferred) or CT scan should also be performed within 28 days prior to randomization and submitted to the imaging core lab. Participants who are found to have untreated brain metastases on the baseline brain scan may not be randomized.

9.1.2.1 Investigator Assessment of Progression

The same method of assessment used at Screening should be used for on-study time points. Brain MRI or CT scans during on-study time points and the follow-up phase are only required in participants with a history of CNS metastases prior to randomization or if clinically indicated for new signs or symptoms that suggest new or worsening CNS metastases.

Post-baseline tumor assessments will be performed at the time points described below until progression assessed by the investigator **and** confirmed by BICR, death, or withdrawal from the study, whichever occurs first.

- First tumor assessment post-baseline should be performed at Week 12 (± 7 days) following randomization. Use same imaging method as was used at screening/baseline.
- Subsequent tumor assessments should occur at every 6 weeks until Week 60. Allowable window for assessments is ± 7 days until Week 60. After Week 60, tumor assessments should occur every 12 weeks (± 14 days) until radiographic progression, assessed by the investigator **and** confirmed by the BICR.
- Additional imaging of potential disease sites should be performed whenever disease progression or a secondary malignancy is suspected. In participants with no history of brain lesions prior to randomization, brain MRI or CT on-study treatment should be obtained if

clinically indicated. Bone imaging during on-study treatment and follow-up periods should be obtained if clinically indicated.

- Tumor assessments are to continue for all randomized participants according to the protocol schedule until radiographic progression has been assessed by the investigator **and** confirmed by the BICR, regardless of whether study drug dosing is delayed, reduced, or discontinued.

The investigator (in consultation with the local radiologist as needed) will complete tumor assessments using RECIST (Response Evaluation Criteria in Solid Tumors) 1.1 criteria on all imaging time points specified in the protocol. Any additional imaging that may demonstrate tumor response or progression, including scans performed at unscheduled time points and/or at an outside institution, should also be collected for the investigator to complete RECIST 1.1 tumor assessments on these images and to submit them for BICR review.

All study treatment decisions will be based on the investigator's assessment of tumor images and not on the BICR assessment.

9.1.2.2 BICR Assessment of Progression

Sites should submit all scans to a BICR on a rolling basis, preferably within 7 days of scan acquisition, throughout the duration of the study. BICR will review scans on a rolling basis and remain blinded to treatment arm and investigator assessment of submitted scans. When progression per RECIST 1.1 criteria is assessed by the investigator, the site will inform the imaging core lab, so that the BICR assessment of progression can be performed. The BICR review will be completed and the results provided to the site within approximately 14 days of receipt of the scans, provided there are no pending imaging queries to the site.

Participants whose progression is not confirmed by the BICR will be required to continue tumor assessments (if clinically feasible) according to the protocol-specified schedule or sooner if clinically indicated until the BICR confirms progression on a subsequent tumor assessment. Also, if participants discontinue treatment without radiographic progression, tumor assessments will continue according to the protocol specified schedule, as noted in Section 2, until progression has been confirmed by BICR.

All study treatment decisions will be based on the investigator's assessment of tumor images and not on the BICR assessment. The BICR assessment of progression is only relevant for determining when tumor assessments for a given participant are no longer required to be submitted to the imaging vendor.

9.1.3 Imaging Restriction and Precautions

Table 9.1.3-1 provides a summary of the alternative methods, acceptable per protocol, in the event of contraindications for use of IV and oral contrast, and or/MRI.

Table 9.1.3-1: Acceptable Imaging Assessment Methods for Different Anatomic Regions

Anatomic Region	Preferred Method	Alternative Methods

Table 9.1.3-1: Acceptable Imaging Assessment Methods for Different Anatomic Regions

Anatomic Region	Preferred Method	Alternative Methods
Chest, abdomen, and pelvis Note: Scan must cover lung apices to diaphragm, diaphragm through entire liver, and to below the pubic symphysis	CT with IV contrast	**For chest:** • CT without contrast can be used only if the participant has a clinical contraindication for iodine-based IV contrast (eg, hypersensitivity, renal insufficiency) **For abdomen and pelvis:** • MRI with gadolinium-based IV contrast is the first alternative method if the participant has a clinical contraindication for iodine-based IV contrast • CT without contrast can be used as the second alternative method only if the participant has a clinical contraindication for both contrast-enhanced CT and MRI.
Brain	MRI with IV contrast	• CT with IV contrast is the first alternative method if IV gadolinium is clinically contraindicated. • MRI without contrast can be used as a second alternative method if a participant has clinical contraindications for both contrast-enhanced CT and MRI
Bone	Bone scintigraphy	PET (18F-fluoride NaF or FDG) and 99m Technetium SPECT

Notes:

1. CT scans must be performed with slices thickness of ≤ 5 mm are required. The reconstruction interval should be equal to slice thickness to avoid gap.
2. The same modality for a given anatomical coverage and the same scanning procedure (most importantly: reconstruction slice thickness, anatomic coverage, use of IV contrast) should be consistent between baseline and all subsequent follow-up scanning. If possible, the same scanner or an equivalent scanners should be used throughout the study.
3. For abdomen and pelvis CT scans, oral contrast is recommended as per institutional standards.
4. MRI should include T1 and T2-weighted sequences with T1-weighted at both pre- and post-contrast.
5. If bone scan shows hotspots indicative of metastases, further investigation with X-ray, CT, or MRI is warranted.
6. All scans generated should be exportable in electronic format (DICOM) to enable secure and rapid electronic transmission to the designated central imaging laboratory.

The use of gadolinium-based contrast agents in participants with acute or chronic renal insufficiency, with a glomerular filtration rate (GFR) less than 30 mL per minute per 1.73m^2 or with any acute renal failure caused by hepatorenal syndrome or perioperative liver transplantation, is not recommended.

If gadolinium is contraindicated, proceed without contrast but reason for not using contrast must be documented.

9.1.4 Outcomes Research Assessments

Participant-reported outcomes will be captured through the use of 2 validated self-reported questionnaires: the National Comprehensive Cancer Network (NCCN) Functional Assessment of Cancer Therapy - Kidney Symptom Index (FKSI-19), and the EuroQoL Group's EQ-5D-3L.

The NCCN FKSI-19 is a 19-item scale that measures tumor specific HrQoL in kidney cancer participants. The FKSI-19 uses 5 Likert-type response categories that range from "not at all" to "very much." Participants are asked to circle the response category that best characterizes their response over the last 7 days on 19 items that include symptoms such as lack of energy, fatigue, appetite, coughing, shortness of breath, pain, nausea, and ability to work.

General health status will be measured using the EQ-5D-3L. The EQ-5D-3L is a standardized instrument for use as a measure of self-reported general health status. The EQ-5D-3L comprises 5 dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety) and a visual analog rating scale (VAS). The utility data generated from EQ-5D-3L is recommended for and commonly used in cost effectiveness analysis.

9.2 Adverse Events

The definitions of an AE or serious adverse event (SAE) can be found in Appendix 3.

AEs will be reported by the participant (or, when appropriate, by a caregiver, surrogate, or the participant's legally authorized representative).

The investigator and any designees are responsible for detecting, documenting, and reporting events that meet the definition of an AE or SAE and remain responsible for following up AEs that are serious, considered related to the study treatment or the study, or that caused the participant to discontinue before completing the study.

Contacts for SAE reporting specified in Appendix 3

Immune-mediated adverse events (IMAEs, imAEs) are AEs consistent with an immune-mediated mechanism or immune-mediated component for which non-inflammatory etiologies (eg, infection or tumor progression) have been ruled out. IMAEs can include events with an alternate etiology which were exacerbated by the induction of autoimmunity. Information supporting the assessment will be collected on the participant's case report form.

9.2.1 Time Period and Frequency for Collecting AE and SAE Information

The collection of nonserious AE information should begin at initiation of study treatment until [the follow-up contact], at the timepoints specified in the Schedule of Activities (Section 2). Nonserious AE information should also be collected from the start of a placebo lead-in period or other observational period intended to establish a baseline status for the participants.

Sections 5.6.1 and 5.6.2 in the Investigator Brochures (IBs) for Nivolumab[29] and Ipilimumab[30] represent the Reference Safety Information (also Appendix K in the IB for cabozantinib[22]) to determine expectedness of serious adverse events for expedited reporting. Following the participant's written consent to participate in the study, all SAEs, whether related or not related to study drug, must be collected, including those thought to be associated with protocol-specified procedures.

All SAEs must be collected that occur during the screening period and within 100 days of the last dose of study treatment. For participants randomized to treatment and never treated with

study drug, SAEs should be collected for 30 days from the date of randomization. If applicable, SAEs must be collected that relate to any later protocol-specified procedure (eg, a follow-up skin biopsy).

The investigator must report any SAE that occurs after these time periods and that is believed to be related to study drug or protocol-specified procedure.

- Medical occurrences that begin before the start of study treatment but after obtaining informed consent will be recorded on the appropriate section of the eCRF section.
- All SAEs will be recorded and reported to Sponsor or designee within 24 hours, as indicated in Appendix 3.
- The investigator will submit any updated SAE data to the sponsor within 24 hours of this being available.

Investigators are not obligated to actively seek AEs or SAEs in former study participants. However, if the investigator learns of any SAE, including a death, at any time after a participant has been discharged from the study, and he/she considers the event reasonably related to the study treatment or study participation, the investigator must promptly notify the sponsor.

The method of evaluating, and assessing causality of AEs and SAEs and the procedures for completing and reporting/transmitting SAE reports are provided in Appendix 3.

9.2.2 Method of Detecting AEs and SAEs

Adverse events can be spontaneously reported or elicited during open-ended questioning, examination, or evaluation of a participant. (In order to prevent reporting bias, participants should not be questioned regarding the specific occurrence of one or more AEs.)

9.2.3 Follow-up of AEs and SAEs

- Nonserious AEs should be followed to resolution or stabilization, or reported as SAEs if they become serious (see Section Appendix 3).
- Follow-up is also required for nonserious AEs that cause interruption or discontinuation of study treatment and for those present at the end of study treatment as appropriate.
- All identified nonserious AEs must be recorded and described on the nonserious AE page of the CRF (paper or electronic). Completion of supplemental CRFs may be requested for AEs and/or laboratory abnormalities that are reported/identified during the course of the study.

After the initial AE/SAE report, the investigator is required to proactively follow each participant at subsequent visits/contacts. All SAEs, and non-serious AEs of special interest (as defined in Section 9.2 will be followed until resolution, until the condition stabilizes, until the event is otherwise explained, or until the participant is lost to follow-up (as defined in Section 8.3).

Further information on follow-up procedures is given in Appendix 3.

9.2.4 Regulatory Reporting Requirements for SAEs

- Prompt notification by the investigator to the Sponsor of SAEs is essential so that legal obligations and ethical responsibilities towards the safety of participants and the safety of a product under clinical investigation are met.

- An investigator who receives an investigator safety report describing SAEs or other specific safety information (eg, summary or listing of SAEs) from the Sponsor will file it along with the Investigator Brochure and will notify the IRB/IEC, if appropriate according to local requirements.

Sponsor or designee will be reporting adverse events to regulatory authorities and ethics committees according to local applicable laws including European Directive 2001/20/EC and FDA Code of Federal Regulations 21 CFR Parts 312 and 320. A SUSAR (Suspected, Unexpected Serious Adverse Reaction) is a subset of SAEs and will be reported to the appropriate regulatory authorities and investigators following local and global guidelines and requirements.

All SAEs must be collected that occur during the screening period and within 100 days of the last dose of study treatment. For participants randomized/assigned to treatment and never treated with study drug, SAEs should be collected for 30 days from the date of randomization.

9.2.5 Pregnancy

If, following initiation of the study treatment, it is subsequently discovered that a participant is pregnant or may have been pregnant at the time of study exposure, including during at least 5 half-lives after product administration, the investigator must immediately notify the BMS Medical Monitor/designee of this event and complete and forward a Pregnancy Surveillance Form to BMS Designee within 24 hours of awareness of the event and in accordance with SAE reporting procedures described in Appendix 3.

In most cases, the study treatment will be permanently discontinued in an appropriate manner (eg, dose tapering if necessary for participant safety). Please call the BMS Medical Monitor within 24 hours of awareness of the pregnancy.

Protocol-required procedures for study discontinuation and follow-up must be performed on the participant.

Follow-up information regarding the course of the pregnancy, including perinatal and neonatal outcome and, where applicable, offspring information must be reported on the Pregnancy Surveillance Form.

Any pregnancy that occurs in a female partner of a male study participant should be reported to Sponsor or designee. In order for Sponsor or designee to collect any pregnancy surveillance information from the female partner, the female partner must sign an informed consent form for disclosure of this information. Information on this pregnancy will be collected on the Pregnancy Surveillance Form.

9.2.6 Laboratory Test Result Abnormalities

The following laboratory test result abnormalities should be captured on the nonserious AE CRF page or SAE Report Form electronic, as appropriate. Paper forms are only intended as a back-up option when the electronic system is not functioning.

- Any laboratory test result that is clinically significant or meets the definition of an SAE

- Any laboratory test result abnormality that required the participant to have study treatment discontinued or interrupted
- Any laboratory test result abnormality that required the participant to receive specific corrective therapy

It is expected that wherever possible, the clinical rather than laboratory term would be used by the reporting investigator (eg, anemia versus low hemoglobin value).

9.2.7 Potential Drug Induced Liver Injury (DILI)

Wherever possible, timely confirmation of initial liver-related laboratory abnormalities should occur prior to the reporting of a potential DILI event. All occurrences of potential DILIs, meeting the defined criteria, must be reported as SAEs (see Section 9.2 and Appendix 3 for reporting details).

Potential drug induced liver injury is defined as:

- AT (ALT or AST) elevation > 3 times upper limit of normal (ULN)
  AND
- Total bilirubin > 2 times ULN, without initial findings of cholestasis (elevated serum alkaline phosphatase),
  AND
- No other immediately apparent possible causes of AT elevation and hyperbilirubinemia, including, but not limited to, viral hepatitis, pre-existing chronic or acute liver disease, or the administration of other drug(s) known to be hepatotoxic.

9.2.8 Other Safety Considerations

Any significant worsening noted during interim or final physical examinations, electrocardiogram, x-ray filming, any other potential safety assessment required or not required by protocol should also be recorded as a nonserious or serious AE, as appropriate, and reported accordingly.

9.3 Overdose

An overdose is defined as the accidental or intentional administration of any dose of a product that is considered both excessive and medically important. All occurrences of overdose must be reported as SAEs (see Section 9.2).

9.4 Safety

Planned time points for all safety assessments are listed in the Schedule of Activities.

9.4.1 Clinical Safety Laboratory Assessments

Investigators must document their review of each laboratory safety report.

Hematology (CBC)
Hemoglobin
Hematocrit
Total leukocyte count, including differential
Platelet count

Prothrombin time (PT)/ International normalized ratio (INR)	
Partial thromboplastin time (PTT)	
Chemistry	
Aspartate aminotransferase (AST) Alanine aminotransferase (ALT) Total bilirubin Alkaline phosphatase Lactate dehydrogenase (MLR) Creatinine Blood Urea Nitrogen (BUN) Glucose Albumin	Sodium (Na) Potassium (K) Chloride (Cl) Calcium (Ca) Corrected calcium (Screening only) Phosphorus (P) Magnesium (Mg) Amylase Lipase
Urinalysis	
Creatinine Protein Urine protein/creatinine ratio (UPCR)	
Serology	
Serum for hepatitis C antibody, HCV RNA, hepatitis B surface antigen (screening only). HIV, if mandated locally. (Sites in Germany, see Appendix 12)	
Other Analyses	
Thyroid stimulating hormone (TSH) with free thyroxine (fT3) and free triiodothyronine (fT4) (screening only); TSH with reflexive fT3 and fT4 during study and follow up	
Pregnancy test (WOCBP only: screening and during study)	
Follicle stimulating hormone (FSH) (screening only for women under 55 years old to confirm menopause as needed)	

9.4.2 Imaging Safety Assessment

Any incidental findings of potential clinical relevance that are not directly associated with the objectives of the protocol should be evaluated and handled by the Study Investigator as per standard medical/clinical judgment.

9.5 Pharmacokinetics and Immunogenicity

Samples for nivolumab, ipilimumab, and cabozantinib PK and nivolumab and ipilimumab immunogenicity assessments will be collected for participants in Arm A (doublet) receiving nivolumab combined with cabozantinib and in Arm B (triplet) receiving nivolumab and ipilimumab combined with cabozantinib as described in Table 9.5-1 and Table 9.5-2, respectively. Treatment assignments will be released to the bioanalytical laboratory in order to minimize unnecessary analysis of samples. All timepoints for nivolumab/ipilimumab sampling are relative to the start of study drug administration for nivolumab in Arms A and B. All timepoints for cabozantinib PK sampling are relative to the start of study drug administration for nivolumab in Arms A and B and should be drawn approximately at least 8 hours after the previous evening cabozantinib dose.

All on-treatment timepoints are intended to align with days on which study drug is administered. If it is known that a dose is going to be delayed, then the predose sample for nivolumab/ipilimumab and cabozantinib, if appropriate, should be collected just prior to the delayed dose. However, if a predose sample for nivolumab/ipilimumab is collected but the dose is subsequently delayed, an additional predose sample should not be collected. Further details of sample collection, processing, and shipment will be provided in the laboratory/procedure manual.

Blood samples should be drawn from a site other than the infusion site (ie, contralateral arm) on days of infusion. If the infusion was interrupted, the interruption details will also be documented on the CRF. Further details of pharmacokinetic sample collection and processing will be provided to the site in the laboratory/procedure manual.

Samples collected from participants for immunogenicity analyses of nivolumab and/or ipilimumab will be evaluated for the development of Anti-Drug Antibody (ADA) for nivolumab and/or ipilimumab by validated immunoassays. Samples may also be analyzed for neutralizing ADA response to nivolumab and/or ipilimumab.

Serum concentration analyses for nivolumab and/or ipilimumab will be performed by validated immunoassay bioanalytical method(s) for nivolumab and ipilimumab. Plasma concentration analyses for cabozantinib will be performed by validated liquid chromatography tandem-mass spectrometry (LC-MS-MS).

In addition, selected serum samples may be analyzed by an exploratory method that measures nivolumab and ipilimumab, or detect anti-drug antibodies for technology exploration purposes; exploratory results will not be reported. The corresponding serum samples designated for either PK, immunogenicity or biomarker assessments may also be used for any of those analyses, if required (eg, insufficient sample volume to complete testing or to follow up on suspected immunogenicity related AE).

Table 9.5-1: Pharmacokinetic and Immunogenicity Sampling Schedule for Arm A (Doublet)

Study Day[a] (1 Cycle = 2 weeks)	Event (Relative to Nivolumab Dosing) Hour	Time (Relative to Start of Nivolumab Infusion) Hours:Min	Pharmacokinetic Blood Sample for Nivolumab for Arm A	Immunogenicity Blood Sample for Nivolumab for Arm A	Pharmacokinetic Blood Sample for Cabozantinib for Arm A
C1/D1	Predose nivo[b]	00:00	X	X	X[c]
C1/D1	EOI-nivo[d]	00:30	X		
C3/D1	Predose nivo	00:00			X[e]
C4/D1	Predose nivo	00:00	X	X	X[e]
C7/D1	Predose nivo	00:00	X	X	X[e]
C15/D1	Predose nivo	00:00	X	X	
C23/D1	Predose nivo	00:00	X	X	
Every 16 weeks from C23 to 2 years	Predose nivo	00:00	X	X	
Follow-up Samples - Approximately 30 and 100 days from the discontinuation of study drug	NA	NA	X	X	

Abbreviations: C=cycle; D=day; EOI=end of infusion; NA=not applicable; nivo=nivolumab

[a] If a participant discontinues study drug treatment during the sampling period, they will move to sampling at the follow-up visits.

[b] Predose: All predose samples for nivolumab should be taken prior to the start of nivolumab infusion.

[c] Though cabozantinib will be dosed in the evening on Day 1 of Cycle 1, the cabozantinib predose sample on Day 1 can be drawn at the same time nivolumab predose sample is drawn.

[d] EOI-nivo : End of Infusion samples for nivolumab should be collected immediately (preferably within 2-5 minutes) prior to the end of infusion. If the end of infusion is delayed, the collection of the EOI samples should be delayed accordingly. EOI samples may not be collected from the same IV access as drug was administered, refer to the laboratory manual for additional instructions.

[e] Cabozantinib is preferably dosed at bedtime. A PK sample for cabozantinib will be drawn at the same time when predose PK samples for nivolumab are drawn as long at the time of the draw for cabozantinib is approximately at least 8 hours after the previous evening dose of cabozantinib.

Table 9.5-2: Pharmacokinetic (PK) and Immunogenicity Sampling Schedule for Arm B (Triplet)

Study Day[a] (C1 to C4 Cycles = 3 weeks then C5 onward Cycles =2 weeks)	Event (Relative to Nivolumab Dosing) Hour	Time (Relative to Start of Nivolumab Infusion) Hours:Min	PK Blood Sample for Nivolumab for Arm B	Immunogenicity Blood Sample for Nivolumab for Arm B	PK Blood Sample for Ipilimumab for Arm B	Immunogenicity Blood Sample for Ipilimumab for Arm B	PK Blood Sample for Cabozantinib for Arm B
C1/D1	Predose nivo[b]	00:00	X	X	X	X	X[c]
C1/D1	EOI-ipi[d]	01:30	X		X		
C2/D1	Predose nivo	00:00					X[e]
C3/D1	Predose nivo	00:00	X	X	X	X	X[e]
C5/D1	Predose nivo	00:00	X	X	X	X	X[e]
C13/D1	Predose nivo	00:00	X	X			
C21/D1	Predose nivo	00:00	X	X			
Every 16 weeks from C21 to 2 years	Predose nivo	00:00	X	X			
Follow-up Samples - Approximately 30 and 100 days from the discontinuation of study drug	NA	NA	X	X	X	X	

Abbreviations: C=cycle; D=day; EOI=end of infusion; ipi=ipilimumab; NA=not applicable; nivo=nivolumab

[a] If a participant discontinues study drug treatment during the sampling period, they will move to sampling at the follow-up visits.

[b] Predose: All predose samples for nivolumab and/or ipilimumab should be taken prior to the start of nivolumab infusion.

[c] Though cabozantinib will be dosed in the evening on Day 1 of Cycle 1, the cabozantinib predose sample on Day 1 can be drawn at the same time nivolumab predose sample is drawn.

[d] EOI-ipi: The end of infusion (EOI) samples for both nivolumab and ipilimumab should be collected immediately (preferably within 2-5 minutes) prior to the end of ipilimumab infusion. If the end of infusion is delayed, the collection of the EOI samples should be delayed accordingly. EOI samples may not be collected from the same IV access as drug was administered, refer to the laboratory manual for additional instructions.

[e] Cabozantinib is preferably dosed at bedtime. A PK sample for cabozantinib should be drawn at the same time when predose PK samples for nivolumab and ipilimumab are drawn as long at the time of the draw for cabozantinib is approximately at least 8 hours after the previous evening dose of cabozanitinib.

9.6 Pharmacodynamics

Refer to Section 9.8.

9.7 Pharmacogenomics

Refer to Section 9.8.6.

9.8 Biomarkers

9.8.1 Additional Research Collection

This protocol will include residual sample storage for additional research (AR).

For All US sites:

Additional research is required for all study participants.

- If the IRB/ethics committees and site agree to the mandatory additional research retention and/or collection, then the study participant must agree to the mandatory additional research as a requirement for inclusion in the study.

For non-US Sites

Additional research is optional for all study participants, except where retention and/or collection is prohibited by local laws or regulations, ethics committees, or institutional requirements.

This collection for additional research is intended to expand the translational R&D capability at Bristol-Myers Squibb, and will support as yet undefined research aims that will advance our understanding of disease and options for treatment. It may also be used to support health authority requests for analysis, and advancement of pharmacodiagnostic development to better target drugs to the right participants. This may also include genetic/genomic exploration aimed at exploring disease pathways, progression and response to treatment etc.

Sample Collection and Storage

All requests for access to samples or data for additional research will be vetted through a diverse committee of the study sponsor's senior leaders in Research and Development (or designee) to ensure the research supports appropriate and well-defined scientific research activities.

- Residual whole blood (DNA), serum, plasma, PBMCs, tumor tissue from whole blood (DNA), serum biomarkers, plasma biomarkers, PBMCs, tumor biopsy collections (see Table 9.8.1-1) will also be retained for additional research purposes

Samples kept for future research will be stored at the BMS Biorepository in Hopewell, NJ, USA or an independent, BMS-approved storage vendor.

- The manager of these samples will ensure they are properly used throughout their usable life and will destroy the samples at the end of the scheduled storage period, no longer than fifteen (15) years after the end of the study or the maximum allowed by applicable law.
- Transfers of samples by research sponsor to third parties will be participant to the recipient's agreement to establish similar storage procedures.

Samples will be stored in a coded fashion, and no researcher will have access to the key. The key is securely held by the Investigator at the clinical site, so there is no direct ability for a researcher to connect a sample to a specific individual.

Further details of sample collection and processing will be provided to the site in the procedure manual.

Table 9.8.1-1: Residual Sample Retention for Additional Research Schedule

Sample Type	Timepoints for which residual samples will be retained
Whole blood (DNA)	All
Serum biomarker	All
Plasma biomarker	All
PBMCs	All
Tumor Biopsy	All

9.8.2 Tissue Specimens

Sufficient tumor tissue specimens, preferably collected within 3 months but no more than 12 months prior to enrollment with an associated pathology report, will be required in the form of a formalin-fixed paraffin-embedded block or 20 unstained slides. A minimum of 10 slides will be acceptable if tumor tissue is limited. In these situations, it is recommended to consult with the protocol team to discuss the specifics of the case. In addition, it is recommended, although optional, for tumor tissue samples to be collected upon progression. These samples may be used for the assessment of markers implicated in resistance.

The tumor tissue samples may be used to assess putative predictive biomarkers of nivolumab, ipilimumab, and cabozantinib efficacy and/or to better characterize the tumor-immune microenvironment. Baseline tumor tissue samples will be evaluated for PD-L1 by IHC during screening by the central lab pathologist, who will determine the number of tumor cells with membranous staining among a minimum of 100 evaluable tumor cells. Participants will be stratified at randomization as either PD-L1 expression $\geq$ 1% versus PD-L1 expression < 1% or indeterminate (ie, sample contains at least 100 evaluable tumor cells but membrane scoring is hampered by high cytoplasmic staining).

Various other markers with potential predictive value for the treatment of RCC with nivolumab, ipilimumab and other immunotherapies are currently under investigation and may be assessed in this study. These tumor tissue biomarkers include, but are not limited to PD-L1, PD-1, PD-L2, TILs or subpopulations of TILs and a Th1 immune mRNA expression signature. Evaluation of MET, AXL, and other markers may be performed by IHC or other methods. Tumor samples may also be used to further characterize the tumor-immune microenvironment, including but not limited to other T cell checkpoint receptors and ligands (eg, Lag-3, Tim-3), and intratumoral immune cell subsets, including T-regulatory cells, myeloid derived suppressor cells, macrophages, natural killer (NK) cells and B cells. These samples may also be used to

investigate the effect of nivolumab, ipilimumab, and cabozantinib on the expression of potentially relevant predictive and/or prognostic RCC biomarkers. Both the baseline tumor sample and the sample collected upon progression may be retrospectively assessed for the expression of other immune related genes, RNAs and/or proteins, or for the presence of immune cell populations using a variety of methodologies inclusive of, but not limited to IHC, RNA Scope, qRT-PCR, RNA seq, genetic mutation detection (including whole exome sequencing) and fluorescent in-situ hybridization (FISH).

9.8.3 Exploratory Serum and Plasma Biomarkers

Blood samples for exploratory serum and plasma biomarker analyses will be drawn at the specified time points. Serum and plasma samples may be assessed by ELISA, seromics, microRNA profiling, circulating tumor DNA measurements, metabolomics and/or other relevant multiplex-based protein assay methods for immune or RCC-related factors that will predict for clinical benefit or correlate with treatment-related adverse events. Potential serum and plasma-based biomarkers to potentially b investigated include, but are not limited to levels of soluble PD-L1, anti-tumor antibodies, cytokines, chemokines, inflammatory factors, angiogenic biomarkers (eg, VEGF-A, soluble VEGF2) NKG2D ligands (eg, soluble MICA), circulating tumor DNA, and microRNAs (such as, but not limited to, miR-513 and miR19b).

9.8.4 Myeloid-Derived Suppressor Cells

Myeloid-derived suppressor cells (MDSCs) are an immune cell population capable of suppressing T cell activation and proliferation. MDSCs will be measured at baseline to assess associations with outcome.

9.8.5 PBMC for Flow Cytometry

Exploratory flow cytometry analysis will be used to assess baseline and on-treatment alterations in the composition/ activation status of cell subsets. Lymphocyte subsets to be assayed may include, but are not limited to, CD4+ and CD8+ subsets (activated, effector/memory, regulatory) and populations of those cells as defined by the expression of activation, exhaustion or signaling markers.

9.8.6 Whole Blood for Genotyping

Whole blood samples for exploratory pharmacogenetic assessment will be collected from all participants. Genomic DNA will be extracted and subsequently assessed for SNPs and other genetic variations in candidate genes that may predispose participants to clinical benefit or adverse events (unless restricted by local requirements). Such genes include, but are not limited to, PD-1, PD-L1, PD-L2 and CTLA-4. Additional use of these data may include correlative analyses aimed at identifying genotypic associations with clinically relevant biomarkers identified by other methodologies described in this section (including whole exome sequencing). Genomic DNA from whole blood will be collected and may be used as a comparator for subjects with tumors examined by whole exome or genome analysis.

9.8.7 Other Assessments

Not applicable.

9.9 Health Economics OR Medical Resource Utilization and Health Economics

Health Economics/Medical Resource Utilization and Health Economics parameters will be evaluated in this study as noted in Section 2.

Medical resource utilization and health economics data, associated with medical encounters, will be collected in the CRF by the investigator and study-site personnel for all participants throughout the study. Protocol-mandated procedures, tests, and encounters are excluded.

The data collected may be used to conduct exploratory economic analyses and will include:

- Number and duration of medical care encounters, including surgeries, and other selected procedures (inpatient and outpatient)
- Duration of hospitalization (total days length of stay, including duration by wards; eg, intensive care unit)
- Number and character of diagnostic and therapeutic tests and procedures
- Outpatient medical encounters and treatments (including physician or emergency room visits, tests and procedures, and medications)].

10. STATISTICAL CONSIDERATIONS

10.1 Sample Size Determination

The sample size of this study accounts for the primary endpoint of progression-free survival (PFS) per BICR in the intermediate/poor risk randomized participants. The primary objective specifies two comparisons: PFS per BICR of Arm A (doublet) versus Arm C (sunitinib) and of Arm B (triplet) versus Arm C (sunitinib). This study is not powered for the comparison of Arm A versus Arm B. The sample size and the timing of the analyses are driven by Arm A versus Arm C comparison in intermediate/poor risk randomized participants. Assuming a 25% screen failure rate, it is expected that approximately 1353 any risk participants will need to be enrolled in order to randomize 1014 any risk participants (338 per arm) in a 1:1:1 ratio. This includes at most 273 favorable risk participants (91 per arm) enrolled to randomize 204 favorable risk participants in a 1:1:1 ratio. The rest of the enrolled participants will provide approximately 810 intermediate/poor risk randomized participants (270 per each arm).

The overall alpha for this study is 0.05 (two-sided). This is split with 0.0245 (two-sided) for each of two comparisons to evaluate PFS after penalizing 0.001 (two-sided) to evaluate ORR since it is planned to have an early assessment of ORR. PFS will be evaluated on approximately 810 intermediate/poor risk randomized participants (270 per each arm) for treatment effect at an alpha of 0.0245 (two-sided, penalized 0.0005 from each 0.025 allocation), with at least 90% power. No interim analysis of PFS is planned. OS will be evaluated for treatment effect at an alpha level of 0.0245 (two-sided) with 75% power, accounting for an interim analyses to assess efficacy at the time of the final PFS analysis. ORR will be analyzed on a descriptive basis and will occupy an administrative adjustment of alpha of 0.001 (0.0005 per pairwise comparison).

Sample Size Justification for Primary PFS Endpoint

The primary endpoint of PFS per BICR of Arm A versus Arm C and of Arm B versus Arm C analyses conducted on approximately 810 intermediate/poor risk randomized participants (270 per each arm) will be triggered by approximately 336 events in Arm A and Arm C (or approximately 456 events in all three arms). The 336 PFS events provide at least 90% power to detect a hazard ratio (HR) of 0.68 for PFS of Arm A versus Arm C or a HR of 0.56 for PFS of Arm B versus Arm C with a type I error of 0.0245 (two-sided) for each test for a total of 0.049. The HR of 0.68 corresponds to a 48% increase in the median PFS, assuming a median PFS of 11.8 months for Arm A and 8 months for Arm C. Similarly, the HR of 0.56 corresponds to a 79% increase in the median PFS, assuming a median PFS of 14.3 months for Arm B and 8 months for Arm C. It is projected that an observed hazard ratio of 0.782 or less, which corresponds to a 2.2 month or greater improvement in median PFS (8 versus 10.2 months), would result in a statistically significant improvement in PFS for the Arm A versus Arm C comparison. Similarly, an observed hazard ratio of 0.691 or less, which corresponds to a 3.6 month or greater improvement in median PFS (8 versus 11.6 months), would result in a statistically significant improvement in PFS for the Arm B versus Arm C comparison. An efficacy comparison of between Arm A and Arm B is not an objective of this study, and this study is not powered for such a comparison. Based on the above assumptions, the presumed PFS HR of 0.83 between Arm A and Arm B can be detected with at most 39% power with the specified study design.

For assessing the secondary objectives of this study, a hierarchical testing procedure will be used so that the overall type I error rate is preserved at 0.05. For both Arm A versus Arm C and Arm B versus Arm C comparisons, the key secondary objectives within each pairwise comparison will be tested in the following hierarchical order only if PFS is significant for that pair:

- PFS per BICR assessment among all randomized participants
- OS among intermediate/poor risk randomized participants
- OS among all randomized participants

Sample Size Computation for Secondary OS Endpoint

The secondary objective of OS in intermediate/poor risk participants specifies two comparisons: OS of Arm A versus Arm C and of Arm B versus Arm C. Among the 810 intermediate/poor risk randomized participants (270 per each arm), approximately 295 events (ie, deaths) in Arm A and Arm C (or approximately 414 in all three arms) with an interim analysis at the time of the PFS analysis (when approximately 65% of total OS events have occurred, ie, 192 events in Arms A and C, or approximately 270 in all three arms) provides at least 75% power to detect a HR of 0.71 for OS of Arm A and Arm C and at least 75% power to detect a HR of 0.65 for OS of Arm B vs Arm C with an overall type 1 error of 0.0245 (two-sided) for each test. The HR of 0.71 corresponds to a 41% increase in the median OS, assuming a median OS of 31 months for Arm A and 22 months for Arm C. Similarly, the HR of 0.65 corresponds to a 55% increase in the median OS, assuming a median OS of 34 months for Arm B and 22 months for Arm C.

The stopping boundaries at interim and final OS analyses will be derived based on the actual number of deaths using O'Brien and Fleming α spending function. For example, with 192 and

295 observed events in Arm A and Arm C at the interim and final analyses, the respective stopping boundaries would be α=0.005 (two-sided) and α=0.023 (two-sided). Similarly, with 122 and 187 observed events in Arm B and Arm C at the interim and final analyses, the respective stopping boundaries would be α=0.005 (two-sided) and α=0.023 (two-sided). However, Arm B will have more events than presumed for these calculations at the time of the PFS and final OS analysis since the timing of the analysis is triggered by the Arm A versus Arm C comparison.

Assuming a constant accrual rate (an average rate of 60 participants /month), the accrual will take approximately 17 months. The final PFS analysis is expected to occur approximately 22 months from FPFV, at which time the interim analysis of OS will also be conducted. The final OS analysis is expected to occur approximately 34 months from FPFV. Table 10.1-1 summarizes the results of these calculations.

Table 10.1-1: Summary of sample size parameters and schedule of analyses

Primary/Secondary Endpoints	**PFS (Primary)**		**OS (Secondary)[a]**	
Primary analysis population	Intermediate/poor risk participants (IMDC score ≥ 1)			
Accrual rate per month for all randomized population	60			
Power	90%		75%	
Comparison (Arm vs Arm)	A vs C	B vs C	A vs C	B vs C
Alpha	0.0245 2-sided	0.0245 2-sided	0.0245 2-sided (0.005 at IA, 0.023 at FA)	0.0245 2-sided (0.005 at IA, 0.023 at FA)
Hypothesized Median Control vs exp (months)	8 vs 11.8	8 vs 14.3	22 vs 31	22 vs 34
Hypothesized Hazard ratio	0.68	0.56	0.71	0.65
Critical Hazard ratio (Observed hazard ratio at which a statistically significant difference would be observed) / Difference in median (months) Corresponding to a minimal clinically significant effect size (FA)[b]	0.782 / 2.2	0.691 / 3.6	0.767 / 6.7	0.717 / 8.7
Critical HR at interim analysis-1(IA) / difference in medians	N/A		0.665 / 11.1	0.6 / 14.7
Target number of event for IA (percentage of target events)	N/A		270 (65%) (192 in A vs C)	
Timing of IA from FPFV (months)	N/A		22	
Accrual Duration (months)	17		17	
Timing of final analysis (FA) from FPFV (months)	22		34	

Table 10.1-1: Summary of sample size parameters and schedule of analyses

Primary/Secondary Endpoints	PFS (Primary)	OS (Secondary)[a]
Sample size	810	810
Target number of events (Event Goal)	456 (336 in A vs C)	414 (295 in A vs C)

[a] OS analyses is subject to significance in hierarchical testing strategy for each pairwise comparison.

[b] The difference in median (months) only applies if the control group median is exactly as hypothesized.

Sample Size Computation for Secondary ORR Endpoint

One of the secondary objectives of the study is to evaluate the ORR per BICR in intermediate and poor-risk participants. The primary analysis of ORR in the intermediate and poor-risk randomized participants will be performed at the time prior to the final PFS analysis when the first 420 participants (140 per arm) will have an approximate 9 months minimum follow-up from 7 months of enrollment. Nine months of minimum follow-up allows for 3 months for responses to occur and an additional 6 months follow-up thereafter intended for sufficient follow-up for ORR, adequate safety follow-up as well as duration of response in this population.

The maximum width of the exact two-sided 95% confidence interval (CI) is approximately 17% when the ORR is expected to be in the 20% to 50% range. Table 10.1-2 summarizes the 95% exact CI when observed ORRs are 20% to 50%, respectively.

Table 10.1-2: 95% Exact CI when Observed ORRs are 20% to 50%

Observed ORR	95% Exact CI
20%	(13.7% - 27.6%)
25%	(18.1% - 33.0%)
30%	(22.6% - 38.3%)
35%	(27.1% - 43.5%)
40%	(31.8% - 48.6%)
45%	(36.6% - 53.6%)
50%	(41.4% - 58.6%)

For example if at least 42 responders are observed among the 140 nivolumab and ipilimumab combination intermediate/poor risk randomized participants (ie, ORR $\geq$ 30%) then the lower bound of the 95% CI is above 22.6%.

10.2 Populations for Analyses

All analyses will be performed using the treatment arm as randomized (intent to treat), with the exception of dosing and safety, for which the treatment arm as received will be used. For purposes of analysis, the following populations are defined in Table 10.2-1.

Table 10.2-1: Populations for Analyses

Population	Description
All Enrolled Participants	All participants who sign informed consent and were registered into the IRT.
All Intermediate/Poor Risk Randomized Participants	All randomized participants with baseline IMDC score ≥ 1 at the time randomization who were randomized to any treatment group. This is the primary efficacy analysis population.
ORR Intermediate/Poor Risk Randomized Participants	The first 420 randomized participants with baseline IMDC score ≥ 1 at the time randomization who were randomized to any treatment group. The analysis of ORR in the intermediate and poor-risk randomized participants will be performed for this population.
All Randomized Participants (any risk participants)	All participants who were randomized to any treatment group. Analysis of demography, protocol deviations, baseline characteristics, secondary efficacy analyses, and outcome research analysis will be performed for this population.
All Treated Participants	All participants who received at least one dose of any study medication. This is the primary population for exposure and safety analyses.
Pharmacokinetic Participants	All participants with available data from randomized participants dosed with nivolumab, ipilimumab, or cabozantinib.
Immunogenicity Participants	All participants with available data from randomized participants dosed with nivolumab or ipilimumab.
Biomarker Participants	All participants with available biomarker data from randomized participants.

10.3 Statistical Analyses

The statistical analysis plan will be developed and finalized before database lock and will describe the selection of participants to be included in the analyses, and procedures for accounting for missing, unused, and spurious data. Below is a summary of planned statistical analyses of the primary and secondary endpoints.

10.3.1 Efficacy Analyses

Table 10.3.1-1: Efficacy Analyses

Endpoint	Statistical Analysis Methods
Primary	The primary objective specifies two comparisons: PFS per BICR of Arm A (doublet) versus Arm C (single agent) and Arm B (triplet) versus Arm C (single agent), in all intermediate/poor risk randomized participants. PFS is defined as the time between the date of randomization and the first date of the documented progression, or death due to any cause, whichever occurs first. Participants who die without a reported progression will be considered to have progressed on the date of their death. Participants who did not progress or die will be censored on the date of their last evaluable tumor assessment. Participants who did not have any on study tumor assessments and did not die will be censored on their date of randomization. Participants who started anti-cancer therapy (including tumor-directed radiotherapy and tumor-directed surgery) without a prior reported progression will be censored on the date of their last evaluable tumor assessment prior to the initiation of first subsequent anti-cancer therapy. The primary formal comparisons of PFS will be conducted using a two-sided 0.0245 stratified log-rank test for each comparison, with IMDC scores, PD-L1 tumor expression, and region at screening per IRT as stratification factors among all intermediate/poor risk

Table 10.3.1-1: Efficacy Analyses

Endpoint	Statistical Analysis Methods
	randomized participants. Median PFS will be estimated via the Kaplan-Meier product limit method. Two-sided 95% CI for the median PFS will be computed for each randomized arm. Kaplan-Meier plots of PFS will be presented. Hazard ratios (HR) and corresponding two-sided 97.55% confidence intervals (CI) will be estimated using a Cox proportional hazards model, with treatment arm as a single covariate, stratified by the stratification factors, corresponding to the comparison of PFS. The totality of PFS results will be presented in a single graphical display that includes Kaplan-Meier curves for all treatment arms, the log-rank p-values for the formal comparisons, the HRs and corresponding CIs, and the median PFS estimates and corresponding CIs. The following supportive analyses of PFS will also be conducted: A stratified multivariate Cox regression model will be used in order to estimate the treatment effect after adjustment for possible imbalances in known or potential prognostic factors. The covariates included in this model will be specified in the statistical analysis plan. PFS using an un-stratified log rank test. The hazard ratio associated with treatment will be presented along with the associated two-sided 97.55% CIs. PFS accounting for missing tumor assessment prior to PFS event (progression or death). This analysis will be performed only if at least 10% of events have missing prior tumor assessment. It will apply the following restriction to the primary definition: If the elapsed time between the PFS event and the last assessment immediately prior to the event is two or more missed visits (more than 12 weeks - 10 days), the participant's PFS will be censored at his/her last tumor assessment prior to the PFS event.
Secondary	The first secondary objective specifies two comparisons: PFS per BICR of Arm A versus Arm C and Arm B versus Arm C, in all randomized participants. PFS will be defined similarly to the primary endpoint except that all randomized participants will be used. At the time of the primary endpoint analysis, the first secondary endpoint will be summarized using the same methods as described above for the primary endpoint analysis. The second secondary endpoint of this specifies two comparisons: OS of Arm A versus Arm C and of Arm B versus Arm C in all intermediate/poor risk randomized participants. OS is defined as the time between the date of randomization and the date of death due to any cause. A participant who has not died will be censored at the last known alive date. At the time of the primary endpoint analysis, there will be an interim analysis of OS, and final analysis of OS is planned to occur 34 months of FPFV. OS will be compared between the treatment arms using a two sided, 0.0245 level log-rank test, stratified using IMDC scores, PD-L1 tumor expression, and region at screening as stratification factors among all randomized participants. A similar analysis as in PFS will be conducted for OS. Hazard ratios (HR) and corresponding two-sided 97.55% confidence intervals (CI) will be estimated using a Cox proportional hazards model, with treatment arm as a single covariate, stratified by the stratification factors, corresponding to the comparison of OS. The third secondary endpoint of this specifies two comparisons: OS of Arm A versus Arm C and of Arm B versus Arm C in all randomized participants. OS is defined as the time between the date of randomization and the date of death due to any cause. A participant who has not died will be censored at the last known alive date. As in the second secondary endpoint, the third secondary endpoint will be summarized at the time of the primary endpoint analysis using the same methods as described above for the first secondary endpoint analysis. The fourth secondary endpoint evaluates ORR per BICR for two different populations: all

Table 10.3.1-1: Efficacy Analyses

Endpoint	Statistical Analysis Methods
	intermediate/poor risk randomized participants and all randomized participants. ORR is defined as the proportion of randomized participants who achieve a best response of complete response (CR) or partial response (PR) using the RECIST 1.1 criteria. Best overall response (BOR) is defined as the best response designation recorded between the date of randomization and the date of objectively documented progression per RECIST 1.1 or the date of subsequent therapy (including tumor-directed radiotherapy and tumor-directed surgery), whichever occurs first. For participants without document progression or subsequent therapy, all available response designations will contribute to the BOR assessment. Duration of response (DOR) is defined as the time between the date of first confirmed documented response (CR or PR) to the date of first documented tumor progression (per RECIST 1.1) or death due to any cause, whichever occurs first. Participants who neither progress nor die will be censored on the date of their last tumor assessment. Responders who started anti-cancer therapy without a prior reported progression will be censored on the date of their last evaluable tumor assessment prior to the initiation of first subsequent anti-cancer therapy. TTR is defined as the time from randomization to the date of the first confirmed documented response (CR or PR), as assessed by BICR. DOR and TTR will be evaluated for responders (CR or PR) only. Tumor assessments are scheduled to be performed at Week 12 following randomization, every 6 weeks for the first 12 months and then every 12 weeks until progression. ORR will be analyzed at the time prior to the final PFS analysis when the first 420 participants (140 per arm) will have an approximate 9 months minimum follow-up from 7 months of enrollment. ORRs and corresponding 95% exact CIs will be calculated using the Clopper Pearson method within each treatment arms. A two sided 95% CI for difference of response rate between Arm A and Arm C arms will also be computed. Similarly, a two sided 95% CI for difference of response rate between Arm B and Arm C arms will also be computed. BOR will be tabulated for each treatment group. An estimate of the response rate and an associated exact two-sided 95% CI (Clopper and Pearson) will be presented, by treatment group. Sensitivity analysis based on investigator-determined ORR may also be performed. DOR and TTR will also be evaluated.
Exploratory	Will be described in the statistical analysis plan finalized before database lock.

10.3.2 Safety Analyses

The safety analysis will be performed in all treated participants.

Table 10.3.2-1: Safety Analyses

Endpoint	Statistical Analysis Methods
Primary	The primary endpoint is related to efficacy and there are no primary safety endpoints.
Secondary	Safety will be analyzed at the time of the primary endpoint analysis. Descriptive statistics of safety will be presented using the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) version 4.0 by treatment arm. All AEs, drug-related AEs, SAEs and drug-related SAEs, imAEs, and select AEs will be tabulated using the worst grade per NCI CTCAE version 4.0 criteria by system organ class and preferred term. On-study lab parameters including hematology, coagulation, chemistry, liver function and renal function will be summarized using worse grade per NCI CTCAE version 4.0 criteria.
Exploratory	Will be described in the statistical analysis plan finalized before database lock.

10.3.3 Other Analyses

Pharmacokinetic, pharmacodynamic, and biomarker exploratory analyses will be described in the statistical analysis plan finalized before database lock. The population pharmacokinetics analysis and pharmacodynamic analyses will be presented separately from the main clinical study report.

Immunogenicity may be reported for ADA positive status (such as persistent positive, other positive, only last sample positive, baseline positive) and ADA negative status, relative to baseline. In addition, presence of neutralizing antibody may be reported, if applicable. Effect of immunogenicity on safety/efficacy and biomarkers and PK may be explored.

10.3.4 Interim Analyses

An interim analysis of ORR will be performed at the time prior to the final PFS analysis when the first 420 participants (140 per arm) will have an approximate 9 months minimum follow-up from 7 months of enrollment. Nine months of minimum follow-up allows for 3 months for responses to occur and an additional 6 months follow-up thereafter intended for sufficient follow-up for ORR, adequate safety follow-up as well as duration of response in this population.

An interim analysis of OS is planned at the time of final PFS analysis and it is expected after observing 192 deaths (approximately 65% of the targeted OS events, or 270 in all three arms) have been observed among intermediate/poor risk participants based on above accrual rate and the exponential distribution in each arm. These formal comparisons of OS will allow for early stopping for superiority, and the boundaries for declaring superiority will be derived based on the actual number of deaths using Lan-DeMets spending function with O'Brien and Fleming type of boundary in EAST version 6. If this interim analysis is performed exactly at 192 deaths, the boundary in terms of statistical significance for declaring superiority would be 0.005 (HR=0.665 with 11.1 months improvement in median OS for the Arm A versus Arm C comparison (22 versus 33.1 months) and HR=0.6 with 14.7 months improvement in median OS for the Arm B versus Arm C comparison (22 versus 36.7 months)). The boundary for declaring superiority in terms of statistical significance for the final analysis after 294 events (or approximately 414 in all three arms) would be 0.023 (HR=0.767 with 6.7 months improvement in median OS for the Arm A versus Arm C comparison (22 versus 28.7 months) and HR=0.717 with 8.7 months improvement in median OS for the Arm B versus Arm C comparison (22 versus 30.7 months)). More details are summarized in Table 10.1-1. An independent statistician external to BMS will perform interim analysis in conjunction with a review by a data monitoring committee.

11. REFERENCES

1 Znaor, Ariana, et al. International Variations and Trends in Renal Cell Carcinoma Incidence and Mortality. Eur Urol 2015; 67(3): 519-530.

2 Sun M, Thuret R, Abdollah F, et al. Age-adjusted incidence, mortality, and survival rates of stage-specific renal cell carcinoma in North America: A trend analysis. Eur Urol 2011; 59: 135–41.

3 Hollingsworth JM, Miller DC, Daignault S, et al. Rising incidence of small renal masses: A need to reassess treatment effect. J Natl Cancer Inst. 2006; 98: 1331–4.

4 Banumathy G, Cairns P. Signaling Pathways in Renal Cell Carcinoma. Cancer Biol Ther 2010; 10(7): 658–664.

5 Heng DY, Xie W, Regan MM, et al. External validation and comparison with other models of the International Metastatic Renal-Cell Carcinoma Database Consortium prognostic model: a population-based study. Lancet Oncol 2013; 14(2): 141-8.

6 NCCN Clinical Practice Guidelines in Oncology; Kidney Cancer. Version 2.2017. Accessed at https://www.nccn.org/professionals/physician_gls/pdf/kidney.pdf on 7-Feb-2017.

7 Cabometyx (cabozantinib) US Prescribing Information. April 2016.

8 Liu XD, Hoang A, Zhou L, Kalra S, Yetil A, Sun M, et al. Resistance to Antiangiogenic Therapy Is Associated with an Immunosuppressive Tumor Microenvironment in Metastatic Renal Cell Carcinoma. Canc Immunol Res. 2015; 10: 1017-1029.

9 Final Study Report CA209016. Bristol-Myers Squibb. Princeton, NJ. December 9, 2016. Document Control No. 930108520.

10 Atkins MB, Plimack ER, Puzanov I, Fishman MN, McDermott DF, Cho DC, et al. Axitinib in Combination With Pembrolizumab in Patients With Advanced Renal Cell Carcinoma: Preliminary Safety and Efficacy Results [poster no. 773PD]. European Society for Medical Oncology (ESMO), October 7–11, 2016, Copenhagen, Denmark.

11 Taylor MH, Dutcus CE, Schmidt E, Bagulho T, Li D, Shumaker R, et al. A Phase 1b Trial of Lenvatinib Plus Pembrolizumab in Patients With Selected Solid Tumors [poster no.776PD]. ESMO Annual Conference; October 7–11, 2016; Copenhagen, Denmark.

12 Apolo AB, Tomita Y, Lee MJ, Lee S, Frosch A, Steinberg SM. Effect of Cabozantinib on immunosuppressive subsets in metastatic urotherlial carcinoma [abstract No. 4501]. 2014. J Clin Oncol. 2014; 32 (5s). 2014 ASCO Annual Meeting.

13 Currana MA, Montalvoa W, Yagitab H, and Allisona JP. PD-1 and CTLA-4 combination blockade expands infiltrating T cells and reduces regulatory T and myeloid cells within B16 melanoma tumors. PNAS 2010; 107: 4275-4280.

[14] Pardoll D. Does the immune system see tumors as foreign or self? Annu Rev Immunol 2003; 21: 807-39.

[15] Zitvogel L, Tesniere A, Kroemer G. Cancer despite immunosurveillance: immunoselection and immunosubversion. Nat Rev Immunol 2006; 6: 715-27.

[16] Dunn GP, Bruce AT, Ikeda H, et al. Cancer immunoediting: from immunosurveillance to tumor escape. Nat Immunol 2002; 3: 991-8.

[17] Greenwald RJ, Freeman GH, Sharpe AH. The B7 family revisited. Annu Rev Immunol 2004; 23: 515-48.

[18] Freeman FJ, Long AJ, Iwai Y, et al. Engagement of the PD-1 immunohibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation. J Exp Med 2000; 192(7): 1027-34.

[19] Sharpe AH, Wherry EJ, Ahmed R, et al. The function of programmed cell death 1 and its ligands in regulating autoimmunity and infection. Nature Immunol 2007; 8: 237-45.

[20] Wolchok JD, Hoos A, O'Day S, et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. Clin Cancer Res 2009; 15: 7412-20.

[21] Postow MA, Harding J, Wolchok JD. Targeting immune checkpoints: Releasing the restraints on anti-tumor immunity for participants with melanoma. Cancer J 2012; 18: 153-159.

[22] Investigator Brochure for XL184. Version 12. Exelixis, Inc. South San Francisco, CA. July 2016.

[23] Castellone MD, Carlomagno F, Salvatore G, Santoro M. Receptor tyrosine kinase inhibitors in thyroid cancer. Best Pract Res Clin Endo-crinol Metab 2008; 22: 1023-38

[24] Timar J, Dome B. Antiangiogenic drugs and tyrosine kinases. Anti-cancer Agents Med Chem 2008; 8 :462-9.

[25] McDermott DF, Drake CG, Sznol M, et al. Survival, Durable Response, and Long-Term Safety in Patients With Previously Treated Advanced Renal Cell Carcinoma Receiving Nivolumab. J Clin Oncol. 2015; 33(18): 2013-20.

[26] Yang JC, Hughes M, Kammula U et al. Ipilimumab (Anti-CTLA4 Antibody) causes regression of metastatic renal cell cancer associated with enteritis and hypophysitis. J Immunother 2007; 30(8): 825-830.

[27] Kwilas AR, Ardiani A, Donahue RN, Aftab DT, Hodge JW. Dual effects of a targeted small molecule inhibitor (cabozantinib) on immune mediated killing of tumor cells and immune tumor microenvironment permissiveness when combined with a cancer vaccine. Transl Med 2014; 12: 294.

[28] Tolaney SM, Ziehr DR, Guo H, Ng MR, Barry WT, Higgins MJ, et al. Phase II and Biomarker Study of Cabozantinib in Metastatic Triple Negative Breast Cancer Patients. Oncologist 2016; 21:1-8.

[29] Nivolumab Investigator Brochure. Version 15. Bristol-Myer Squibb Research and Development. Princeton, NJ. June 2016. Document Control No. 930038243.

[30] Ipilimumab Investigator Brochure. Version 19. Bristol-Myer Squibb Research and Development. Princeton, NJ. March 2016. Document Control No. 930017531.

[31] Escudier B, Eisen T, et al. Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2012; 23: 65-71.

[32] Motzer RJ, Hutson TE, Tomczak P, et al. Overall survival and updated results for sunitinib compared with interferon alfa in patients with metastatic renal cell carcinoma. J Clin Oncol 2009; 27: 3584-90.

[33] Motzer RJ, Hutson TE, Cella D, et al. Pazopanib versus sunitinib in metastatic renal-cell carcinoma. N Engl J Med 2013; 369:722-731.

[34] Sutent (sunitinib malate) US Prescribing Information. April 2014.

[35] Motzer RJ, Rini BI, McDermott DF, et al. Nivolumab for Metastatic Renal Cell Carcinoma: Results of a Randomized Phase II Trial. J Clin Oncol 2015; 33(13): 1430-7.

[36] Topalian SL, Hodi FS, Brahmer JR, et al. Safety, Activity, and Immune Correlates of Anti-PD-1 Antibody in Cancer. N Engl J Med. 2012. 366(26): 2443-54.

[37] Motzer RJ, Rini BI, McDermott DF, et al. Nivolumab for Metastatic Renal Cell Carcinoma: Results of a Randomized Phase II Trial. J Clin Oncol 2015; 33(13): 1430-7.

[38] Choueiri TK, Escudier BJ, Kim JJ, et al. LJ. Immunomodulatory activity of nivolumab in previously treated and untreated metastatic renal cell carcinoma (mRCC): Biomarker-based results from a randomized clinical trial. J Clin Oncol 2014; 32:5s (suppl; abstr 5012).

[39] A Phase 2 Open-label Study of Single Agent Ipilimumab for the Treatment of IL-2 Refractory or IL-2 Ineligible Participants with Stage IV Renal Cancer; NCT00057889; Final Clinical Study Report 930025954 v1.0 (28-January-2008).

[40] A Phase 2 Open-label Study of Single Agent Ipilimumab for the Treatment of IL-2 Refractory or IL-2 Ineligible Participants with Stage IV Renal Cancer; NCT00057889; Final Clinical Study Report 930025954 v1.0 (28-January-2008).

[41] Choueiri TK, Escudier B, Powles T, Mainwaring PN, Rini BI, Donskov F, et al, for the METEOR Investigators. Cabozantinib versus Everolimus in Advanced Renal-Cell Carcinoma. N Engl J Med 2015; 373: 1814-23.

[42] Choueiri TK, Cabozantinib versus everolimus in advanced renal cell carcinoma (METEOR): final results from a randomised, open-label, phase 3 trial. Lancet Oncol 2016; 17(7): 917-27.

[43] Choueiri TK, Halabi S, Sanford BL et al. Cabozantinib Versus Sunitinib As Initial Targeted Therapy for Participants With Metastatic Renal Cell Carcinoma of Poor or Intermediate Risk: The Alliance A031203 CABOSUN Trial J Clin Oncol 2017; 35(6):591-597.

[44] Apolo AB, Mortazavi A, Stein MN, Pal SK, Davarpanah NN, Parnes HL, et al. A Phase I/Expansion Study of Cabozantinib Plus Nivolumab (CaboNivo) Alone or With Ipilimumab (CaboNivoIpi) in Patients With Metastatic Urothelial Carcinoma and other Genitourinary Tumors. ASCO Genitourinary Cancers Symposium, Orlando, Florida. February, 2017.

[45] Heng DYC, Xie W, Regan MM, Harshman LC, Bjarnason GA, Vaishampayan UN, et al. External validation and comparison with other models of the International Metastatic Renal-Cell Carcinoma Database Consortium prognostic model: a population-based study. Lancet Oncol 2013; 14: 141-48.

[46] Brahmer J, et al. Nivolumab versus Docetaxel in Advanced Squamous-Cell Non–Small-Cell Lung Cancer. N Engl J Med 2015; 373: 123-35.

[47] Motzer R J, et al. Nivolumab versus everolimus in advanced renal-cell carcinoma. N Engl J Med 2015; 373:1803-13.

[48] Schadendorf D, et al, 2016 Annual Meeting of the European Association of Dermato-Oncology; August 31-September 3, 2016; Vienna, Austria.

[49] Schadendorf D. et al. Pooled Analysis of Long-Term Survival Data From Phase II and Phase III Trials of Ipilimumab in Unresectable or Metastatic Melanoma. J Clin Oncol 2015; 33(17):1889-94.

[50] Bennouna J, Sastre J, Arnold D, et al. Continuation of bevacizumab after first progression in metastatic colorectal cancer (ML18147): a randomised phase 3 trial. Lancet Oncol. 2013; 14:29-37.

[51] Takeda M, Yamanaka T, Seto T, et al. Bevacizumab beyond disease progression after first-line treatment with bevacizumab plus chemotherapy in advanced nonsquamous non–small cell lung cancer (West Japan Oncology Group 5910L): an open-label, randomized, phase 2 trial. Cancer 2016; 122:1050-1059.

12 APPENDICES

APPENDIX 1 ABBREVIATIONS AND TRADEMARKS

Term	Definition
AE	adverse event
ACLS	advanced cardiac life support
AHA	alpha hydroxy acid
AI	accumulation index
AIDS	acquired immunodeficiency syndrome
AJCC	American Joint Committee on Cancer
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANC	absolute neutrophil count
ANOVA	analysis of variance
APC	antigen-presenting cell
aPTT	activated partial thromboplastin time
AST	aspartate aminotransferase
AT	aminotransaminases
AUC	area under the concentration-time curve
AUC(INF)	area under the concentration-time curve from time zero extrapolated to infinite time
AUC(0-T)	area under the concentration-time curve from time zero to the time of the last quantifiable concentration
AUC(TAU)	area under the concentration-time curve in one dosing interval
A-V	atrioventricular
AXL	member of the TAM (Tyro-3, Axl, Mer) receptor tyrosine kinases (RTK) subfamily
β-HCG	beta-human chorionic gonadotrophin
Bcl-xL	anti-apoptotic member of the B-cell lymphoma 2 (BCL-2) protein family
BICR	blinded independent central review
BID, bid	bis in die, twice daily
BLQ	below limit of quantification
BMI	body mass index
BMS	Bristol-Myers Squibb

Term	Definition
BOR	best overall response
BP	blood pressure
BTLA	B and T lymphocyte attenuator
BUN	blood urea nitrogen
C	Celsius
C	cycle
C12	concentration at 12 hours
C24	concentration at 24 hours
Ca, Ca^{++}	calcium
Cavg	average concentration
Cavgss	average concentration at steady state
CBC	complete blood count
CD	cluster of differentiation
CD3, CD8, CD14, CD28	cluster of differentiation 3, 8, 14, 28
CD137	member of the tumor necrosis factor (TNF) receptor family. Alternative names are TNF receptor superfamily member 9 (TNFRSF9), 4-1BB and induced by lymphocyte activation (ILA)
Cexpected-tau	expected concentration in a dosing interval
CFR	Code of Federal Regulations
CHF	congestive heart failure
CI	confidence interval
Cl, $C1^{-}$	chloride
CLcr	creatinine clearance
CLT	total body clearance
CLT/F (or CLT)	apparent total body clearance
CLT/F/fu or CLT/fu	Apparent clearance of free drug or clearance of free if (if IV)
cm	centimeter
Cmax, CMAX	maximum observed concentration

Term	Definition
c-MET	tyrosine-protein kinase mesenchymal-epithelial transition (MET) or hepatocyte growth factor receptor, is a protein that in humans is encoded by the MET gene
CMH	Cochran-Mantel-Haenzel
Cmin, CMIN	minimum observed concentration
CMV	cytomegalovirus
CNS	central nervous system
CONSORT	Consolidated standards of reporting trials
CR	complete response
CRC	Clinical Research Center
CRF	Case Report Form, paper or electronic
CT	computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
CTLA-4	cytotoxic T lymphocyte antigen-4
Ctrough	Trough observed plasma concentration
CV	coefficient of variation
CVA	cerebrovascular accident
CYP	cytochrome p-450
D	day
DBP	diastolic blood pressure
D/C	discontinue
DILI	drug induced liver injury
dL	deciliter
DMC	data monitoring committee
DNA	deoxyribonucleic acid
DOR	duration of response
DSM IV	Diagnostic and Statistical Manual of Mental Disorders (4^{th} Edition)
DVT	deep vein thrombosis
EA	extent of absorption
EC50	half maximal effective concentration
ECG	electrocardiogram

Term	Definition
eCRF	Electronic Case Report Form
EDC	Electronic Data Capture
EEG	electroencephalogram
e.g., eg	exempli gratia (for example)
eGFR	estimated glomerular filtration rate
ELISA	enzyme-linked immunosorbent assay
EOI	end of infusion
EQ-5D-3L	EuroQoL Group's instrument to measure general health status
ESR	Expedited Safety Report
F	bioavailability
FA	final analysis
FDA	Food and Drug Administration
FDG-PET	fludeoxyglucose- positron emission tomography
FFPE	formalin-fixed, paraffin-embedded
FISH	fluorescent in-situ hybridization
FKSI-19	Functional Assessment of Cancer Therapy - Kidney Symptom Index
FLT-3	Fms-related tyrosine kinase 3
FSH	follicle stimulating hormone
fT3, fT4	free thyroxine (fT3), free triiodothyronine (fT4)
FU	follow up
g	gram
GC	gas chromatography
GCP	Good Clinical Practice
GGT	gamma-glutamyl transferase
GFR	glomerular filtration rate
GI	gastrointestinal
h	hour
HA	health authorities
HBsAg	hepatitis B surface antigen
HBV	hepatitis B virus

Term	Definition
hCG, HCG	human chorionic gonadotropin
HCO_3^-	bicarbonate
HCV	hepatitis C virus
HFS	hand foot syndrome
HIV	Human Immunodeficiency Virus
HR	heart rate, hazard ratio
HrQoL	health-related quality of life
HRT	hormone replacement therapy
ICD	International Classification of Diseases
IC50	half maximal inhibitory concentration
ICH	International Conference on Harmonisation
ICOS	inducible co-stimulator
i.e., ie	id est (that is)
IEC	Independent Ethics Committee
IFN-γ	interferon-γ
IFN-alpha	interferon alphas
IgG1	immunoglobulin G1
IHC	Immunohistochemistry
IL	interleukin
IL-2	interleukin-2
IMAEs, imAEs	immune-mediated adverse events
IMDC	International Metastatic Renal Cell Carcinoma Database Consortium
IMP	investigational medicinal products
IND	Investigational New Drug Exemption
INR	international normalized ratio
IP	investigational product
ipi	ipilimumab
iRAEs	immune-related adverse events
IRB	Institutional Review Board
IRC	independent radiologic review committee

Term	Definition
IRT	Interactive Response Technology
IU	International Unit
IV	intravenous
K	slope of the terminal phase of the log concentration-time curve
K_3EDTA	potassium ethylenediaminetetraacetic acid
K, K^+	potassium
kg	kilogram
KIT	platelet-derived growth factor receptors (PDGFRs)
KPS	Karnofsky Performance Status
L	liter
LAM	Lactation amenorrhea method
LC	liquid chromatography
LC-MS-MS	liquid chromatography tandem-mass spectrometry
LDH	lactate dehydrogenase
LFT	liver function test
LINAC	linear accelerator
LLN	lower limit of normal
LMWH	low molecular weight heparin
ln	natural logarithm
MDSCs	myeloid-derived suppressor cells
MER	proto-oncogene tyrosine-protein kinase MER is an enzyme that in humans is encoded by the MERTK gene
MET	mesenchymal-epithelial transition factor, a tyrosine kinase receptor
mg	milligram
Mg, Mg^{++}	magnesium
MHC	major histocompatibility complex
min	minute
mL	milliliter
MLR	mixed lymphocyte reaction
mmHg	millimeters of mercury

Term	Definition
mOS	median overall survival
mRCC	metastatic renal cell carcinoma
MRI	magnetic resonance imaging
MS	mass spectrometry
MSKCC	Memorial Sloan Kettering Cancer Center
MTD	maximum tolerated dose
mTOR	mammalian target of rapamycin
mUC	metastatic urothelial carcinoma
MUGA	multigated acquisition scan
mWHO	modified World Health Organization
μg	microgram
N	number of subjects or observations
N/A, NA	not applicable
Na, Na^+	sodium
NaF	Sodium fluoride
NCCN	National Comprehensive Cancer Network
NCI	National Cancer Institute
NE	not evaulable
ng	nanogram
NIMP	non-investigational medicinal products
nivo	nivolumab
NK	natural killer
NKG2D	encoded by KLRK1 gene which is located in the NK-gene complex (NKC)
NR	not reached
NSAID	nonsteroidal anti-inflammatory drug
NSCLC	non-small cell lung cancer
ORR	objective response rate
OS	overall survival
P	phosporous

Term	Definition
PCR	polyermase chain reaction
PD	progressive disease, disease progression
PD	pharmacodynamics
PD-1	programmed death-1
PD-L1	programmed death-ligand 1
PE	pulmonary embolism
PET	positron emission tomography
PFS	progression-free survival
PK	pharmacokinetics
PO	per os (by mouth route of administration)
PR	partial response
PT	prothrombin time
PTT	partial thromboplastin time
Q2W, Q3W	every 2 weeks, every 3 weeks
QC	quality control
QD, qd	quaque die, once daily
qRT-PCR	quantitative real time polyermase chain reaction
QTcF	Fridericia corrected QT
R^2	coefficient of determination
RBC	red blood cell
RCC	renal cell carcinoma
RECIST	Response Evaluation Criteria in Solid Tumors
RET	proto-oncogene encodes a receptor tyrosine kinase for members of the glial cell line-derived neurotrophic factor family of extracellular signalling molecules
RNA	ribonucleic acid
ROS1	proto-oncogene tyrosine-protein kinase ROS is an enzyme that in humans is encoded by the ROS1 gene
ROW	rest of the world
RR	respiratory rate
RS	radiosurgery

Term	Definition
RTK	receptor tyrosine kinases
SAE	serious adverse event
SBP	systolic blood pressure
SCC	squamous cell carcinoma
SD	standard deviation, stable disease
SNP	single-nucleotide polymorphism
SOP	Standard Operating Procedures
Subj	subject
SVC	superior vena cava
t	temperature
T	time
TAMs	tumor-assisted macrophages
TAO	Trial Access Online, the BMS implementation of an EDC capability
TCR	T-cell receptor
T-HALF	Half life
TID, tid	ter in die, three times a day
TIE-2	tunica interna endothelial cell kinase 2
TIL	tumour-infiltrating lymphocyte,
TKIs	tyrosine kinase inhibitors
Tmax, TMAX	time of maximum observed concentration
Tregs	regulatory T cells
TRKB	tropomyosin receptor kinase B (TrkB), also known as tyrosine receptor kinase B
TSH	thyroid stimulating hormone
TTR	time to response
TYRO3	tyrosine-protein kinase receptor TYRO3 is an enzyme that in humans is encoded by the TYRO3 gene
ULN	upper limit of normal
UPCR	urine protein to creatinine ratio
UV	ultraviolet
VAS	visual analog rating scale

Term	Definition
VEGF	vascular endothelial growth factor
VEGFR2	vascular endothelial growth factor receptor 2
Vss/F (or Vss)	apparent volume of distribution at steady state
Vz	volume of distribution of terminal phase (if IV and if multi-exponential decline)
W	washout
WBC	white blood cell
WHO	World Health Organization
wks	weeks
WOCBP	women of childbearing potential
WNOCBP	women **not** of childbearing potential
x g	times gravity
XL184	cabozantinib

APPENDIX 2 STUDY GOVERANCE CONSIDERATIONS

The term 'Participant' is used in the protocol to refer to a person who has consented to participate in the clinical research study. The term 'Subject' used in the eCRF is intended to refer to a person (Participant) who has consented to participate in the clinical research study.

REGULATORY AND ETHICAL CONSIDERATIONS

GOOD CLINICAL PRACTICE

This study will be conducted in accordance with:

- Good Clinical Practice (GCP),
- as defined by the International Council on Harmonisation (ICH)
- in accordance with the ethical principles underlying European Union Directive 2001/20/EC
- United States Code of Federal Regulations, Title 21, Part 50 (21CFR50)
- applicable local requirements.

The study will be conducted in compliance with the protocol. The protocol and any amendments and the participant informed consent will receive approval/favorable opinion by Institutional Review Board/Independent Ethics Committee (IRB/IEC), and regulatory authorities according to applicable local regulations prior to initiation of the study.

All potential serious breaches must be reported to Sponsor or designee immediately. A serious breach is a breach of the conditions and principles of GCP in connection with the study or the protocol, which is likely to affect, to a significant degree, the safety or physical or mental integrity of the subjects of the study or the scientific value of the study.

Personnel involved in conducting this study will be qualified by education, training, and experience to perform their respective tasks.

This study will not use the services of study personnel where sanctions have been invoked or where there has been scientific misconduct or fraud (eg, loss of medical licensure, debarment).

INSTITUTIONAL REVIEW BOARD/INDEPENDENT ETHICS COMMITTEE

Before study initiation, the investigator must have written and dated approval/favorable opinion from the IRB/IEC for the protocol, consent form, participant recruitment materials (eg, advertisements), and any other written information to be provided to subjects. The investigator or BMS should also provide the IRB/IEC with a copy of the Investigator Brochure or product labeling information to be provided to subjects and any updates.

The investigator, Sponsor or designee should provide the IRB/IEC with reports, updates and other information (eg, expedited safety reports, amendments, and administrative letters) according to regulatory requirements or institution procedures.

COMPLIANCE WITH THE PROTOCOL AND PROTOCOL REVISIONS

The investigator should not implement any deviation or change to the protocol without prior review and documented approval/favorable opinion of an amendment from the IRB/IEC (and if applicable, also by local health authority) except where necessary to eliminate an immediate hazard(s) to study subjects.

If a deviation or change to a protocol is implemented to eliminate an immediate hazard(s) prior to obtaining relevant approval/favorable opinion(s) the deviation or change will be submitted, as soon as possible to:

- IRB/IEC for
- Regulatory Authority(ies), if applicable by local regulations (per national requirements)

Documentation of approval/favorable opinion signed by the chairperson or designee of the IRB(s)/IEC(s) and if applicable, also by local health authority must be sent to BMS.

If an amendment substantially alters the study design or increases the potential risk to the participant: (1) the consent form must be revised and submitted to the IRB(s)/IEC(s) for review and approval/favorable opinion; (2) the revised form must be used to obtain consent from subjects currently enrolled in the study if they are affected by the amendment; and (3) the new form must be used to obtain consent from new subjects prior to enrollment.

If the revision is done via an administrative letter, investigators must inform their IRB(s)/IEC(s).

FINANCIAL DISCLOSURE

Investigators and sub-Investigators will provide the Sponsor with sufficient, accurate financial information in accordance with local regulations to allow the Sponsor to submit complete and accurate financial certification or disclosure statements to the appropriate health authorities. Investigators are responsible for providing information on financial interests during the course of the study and for 1 year after completion of the study.

INFORMED CONSENT PROCESS

Investigators must ensure that subjects are clearly and fully informed about the purpose, potential risks, and other critical issues regarding clinical studies in which they volunteer to participate.

In situations where consent cannot be given to subjects, their legally acceptable representatives (as per country guidelines) are clearly and fully informed about the purpose, potential risks, and other critical issues regarding clinical studies in which the participant volunteers to participate.

Sponsor or designee will provide the investigator with an appropriate (ie, Global or Local) sample informed consent form which will include all elements required by ICH, GCP and applicable regulatory requirements. The sample informed consent form will adhere to the ethical principles that have their origin in the Declaration of Helsinki.

Investigators must:

- Provide a copy of the consent form and written information about the study in the language in which the participant is most proficient prior to clinical study participation. The language must be non-technical and easily understood.
- Allow time necessary for participant or participant's legally acceptable representative to inquire about the details of the study.
- Obtain an informed consent signed and personally dated by the participant or the participant's legally acceptable representative and by the person who conducted the informed consent discussion.
- Obtain the IRB/IEC's written approval/favorable opinion of the written informed consent form and any other information to be provided to the subjects, prior to the beginning of the study, and after any revisions are completed for new information.

If informed consent is initially given by a participant's legally acceptable representative or legal guardian, and the participant subsequently becomes capable of making and communicating his or her informed consent during the study, consent must additionally be obtained from the participant.

Revise the informed consent whenever important new information becomes available that is relevant to the participant's consent. The investigator, or a person designated by the investigator, should fully inform the participant or the participant's legally acceptable representative or legal guardian, of all pertinent aspects of the study and of any new information relevant to the participant's willingness to continue participation in the study. This communication should be documented.

The confidentiality of records that could identify subjects must be protected, respecting the privacy and confidentiality rules applicable to regulatory requirements, the subjects' signed ICF and, in the US, the subjects' signed HIPAA Authorization.

The consent form must also include a statement that BMS and regulatory authorities have direct access to participant records.

Subjects unable to give their written consent (eg, stroke or subjects with or severe dementia) may only be enrolled in the study with the consent of a legally acceptable representative. The participant must also be informed about the nature of the study to the extent compatible with his or her understanding, and should this participant become capable, he or she should personally sign and date the consent form as soon as possible. The explicit wish of a participant who is unable to give his or her written consent, but who is capable of forming an opinion and assessing information to refuse participation in, or to be withdrawn from, the clinical study at any time should be considered by the investigator.

The rights, safety, and well-being of the study subjects are the most important considerations and should prevail over interests of science and society.

SOURCE DOCUMENTS

The Investigator is responsible for ensuring that the source data are accurate, legible, contemporaneous, original and attributable, whether the data are hand-written on paper or entered

electronically. If source data are created (first entered), modified, maintained, archived, retrieved, or transmitted electronically via computerized systems (and/or any other kind of electronic devices) as part of regulated clinical trial activities, such systems must be compliant with all applicable laws and regulations governing use of electronic records and/or electronic signatures. Such systems may include, but are not limited to, electronic medical/health records (EMRs/EHRs), adverse event tracking/reporting, protocol required assessments, and/or drug accountability records).

When paper records from such systems are used in place of electronic format to perform regulated activities, such paper records should be certified copies. A certified copy consists of a copy of original information that has been verified, as indicated by a dated signature, as an exact copy having all of the same attributes and information as the original.

STUDY TREATMENT RECORDS

Records for study treatments (whether supplied by BMS, its vendors, or the site) must substantiate study treatment integrity and traceability from receipt, preparation, administration, and through destruction or return. Records must be made available for review at the request of BMS/designee or a Health Authority.

If	Then
Supplied by BMS (or its vendors):	Records or logs must comply with applicable regulations and guidelines and should include: • amount received and placed in storage area • amount currently in storage area • label identification number or batch number • amount dispensed to and returned by each participant, including unique participant identifiers • amount transferred to another area/site for dispensing or storage • nonstudy disposition (eg, lost, wasted) • amount destroyed at study site, if applicable • amount returned to BMS • retain samples for bioavailability/bioequivalence, if applicable • dates and initials of person responsible for Investigational Product dispensing/accountability, as per the Delegation of Authority Form.
Sourced by site, and not supplied by BMS or its vendors (examples include IP sourced from the sites stock or commercial supply, or a specialty pharmacy)	The investigator or designee accepts responsibility for documenting traceability and study drug integrity in accordance with requirements applicable under law and the SOPs/standards of the sourcing pharmacy.

If	Then
	These records should include: • label identification number or batch number • amount dispensed to and returned by each participant, including unique participant identifiers • dates and initials of person responsible for Investigational Product dispensing/accountability, as per the Delegation of Authority Form.

BMS or designee will provide forms to facilitate inventory control if the investigational site does not have an established system that meets these requirements.

CASE REPORT FORMS

An investigator is required to prepare and maintain adequate and accurate case histories designed to record all observations and other data pertinent to the investigation on each individual treated or entered as a control in the investigation. Data that are derived from source documents and reported on the CRF must be consistent with the source documents or the discrepancies must be explained. Additional clinical information may be collected and analyzed in an effort to enhance understanding of product safety. CRFs may be requested for AEs and/or laboratory abnormalities that are reported or identified during the course of the study.

For sites using the Sponsor or designee electronic data capture tool, electronic CRFs will be prepared for all data collection fields except for fields specific to SAEs and pregnancy, which will be reported on the electronic SAE form and Pregnancy Surveillance form, respectively. If electronic SAE form is not available, a paper SAE form can be used. Spaces may be left blank only in those circumstances permitted by study-specific CRF completion guidelines provided by Sponsor or designee.

The confidentiality of records that could identify subjects must be protected, respecting the privacy and confidentiality rules in accordance with the applicable regulatory requirement(s).

The investigator will maintain a signature sheet to document signatures and initials of all persons authorized to make entries and/or corrections on CRFs.

The completed CRF, SAE/pregnancy CRFs, must be promptly reviewed, signed, and dated by the investigator or qualified physician who is a subinvestigator and who is delegated this task on the Delegation of Authority Form. Subinvestigators in Japan may not be delegated the CRF approval task For electronic CRFs, review and approval/signature is completed electronically through the BMS electronic data capture tool. The investigator must retain a copy of the CRFs including records of the changes and corrections.

Each individual electronically signing electronic CRFs must meet Sponsor or designee training requirements and must only access the BMS electronic data capture tool using the unique user account provided by Sponsor or designee. User accounts are not to be shared or reassigned to other individuals

MONITORING

Sponsor or designee representatives will review data centrally to identify potential issues to determine a schedule of on-site visits for targeted review of study records.

Representatives of BMS must be allowed to visit all study site locations periodically to assess the data quality and study integrity. On site they will review study records and directly compare them with source documents, discuss the conduct of the study with the investigator, and verify that the facilities remain acceptable. Certain CRF pages and/or electronic files may serve as the source documents:

In addition, the study may be evaluated by Sponsor or designee internal auditors and government inspectors who must be allowed access to CRFs, source documents, other study files, and study facilities. BMS audit reports will be kept confidential.

The investigator must notify BMS promptly of any inspections scheduled by regulatory authorities, and promptly forward copies of inspection reports to Sponsor or designee.

RECORDS RETENTION

The investigator (or head of the study site in Japan) must retain all study records and source documents for the maximum period required by applicable regulations and guidelines, or institution procedures, or for the period specified by BMS or designee, whichever is longer. The investigator (or head of the study site in Japan) must contact BMS prior to destroying any records associated with the study.

BMS or designee will notify the investigator (or head of the study site in Japan) when the study records are no longer needed.

If the investigator withdraws from the study (eg, relocation, retirement), the records shall be transferred to a mutually agreed upon designee (eg, another investigator, study site, IRB). Notice of such transfer will be given in writing to BMS or designee.

RETURN OF STUDY TREATMENT

For this study, study treatments (those supplied by BMS, a vendor or sourced by the investigator) such as partially used study treatment containers, vials and syringes may be destroyed on site.

If..	Then
Study treatments supplied by BMS (including its vendors	Any unused study treatments supplied by BMS can only be destroyed after being inspected and reconciled by the responsible Study Monitor unless study treatments containers must be immediately destroyed as required for safety, or to meet local regulations (eg, cytotoxics or biologics).

If..	Then
	If study treatments will be returned, the return will be arranged by the responsible Study Monitor.
Study treatments sourced by site, not supplied by BMS (or its vendors) (examples include study treatments sourced from the sites stock or commercial supply, or a specialty pharmacy)	It is the investigator's or designee's responsibility to dispose of all containers according to the institutional guidelines and procedures.

It is the investigator's or designee's responsibility to arrange for disposal, provided that procedures for proper disposal have been established according to applicable federal, state, local, and institutional guidelines and procedures, and provided that appropriate records of disposal are kept. The following minimal standards must be met:

- On-site disposal practices must not expose humans to risks from the drug.
- On-site disposal practices and procedures are in agreement with applicable laws and regulations, including any special requirements for controlled or hazardous substances.
- Written procedures for on-site disposal are available and followed. The procedures must be filed with the site's SOPs and a copy provided to BMS upon request.
- Records are maintained that allow for traceability of each container, including the date disposed of, quantity disposed, and identification of the person disposing the containers. The method of disposal, ie, incinerator, licensed sanitary landfill, or licensed waste disposal vendor must be documented.
- Accountability and disposal records are complete, up-to-date, and available for the Monitor to review throughout the clinical trial period.

It is the investigator's or designee's responsibility to arrange for disposal of all empty containers.

If conditions for destruction cannot be met the responsible Study Monitor will make arrangements for return of study treatments provided by BMS (or its vendors). Destruction of non- study treatments sourced by the site, not supplied by BMS, is solely the responsibility of the investigator or designee.

CLINICAL STUDY REPORT AND PUBLICATIONS

A Signatory Investigator must be selected to sign the clinical study report.

For this protocol, the Signatory Investigator will be selected as appropriate based on the following criteria:

- National Coordinating Investigator
- Study Steering Committee chair or their designee

- Participant recruitment (eg, among the top quartile of enrollers)
- Involvement in trial design
- Regional representation (eg, among top quartile of enrollers from a specified region or country)
- Other criteria (as determined by the study team)

The data collected during this study are confidential and proprietary to Sponsor or designee. Any publications or abstracts arising from this study must adhere to the publication requirements set forth in the clinical trial agreement (CTA) governing [Study site or Investigator] participation in the study. These requirements include, but are not limited to, submitting proposed publications to Sponsor or designee at the earliest practicable time prior to submission or presentation and otherwise within the time period set forth in the CTA.

APPENDIX 3 ADVERSE EVENTS AND SERIOUS ADVERSE EVENTS: DEFINITIONS AND PROCEDURES FOR RECORDING, EVALUATING, FOLLOW UP AND REPORTING

ADVERSE EVENTS

Adverse Event Definition:
An Adverse Event (AE) is defined as any new untoward medical occurrence or worsening of a preexisting medical condition in a clinical investigation participant administered study drug and that does not necessarily have a causal relationship with this treatment.
An AE can therefore be any unfavorable and unintended sign (such as an abnormal laboratory finding), symptom, or disease temporally associated with the use of study drug, whether or not considered related to the study drug.

SERIOUS ADVERSE EVENTS

Serious Adverse Event (SAE) is defined as any untoward medical occurrence that, at any dose:
Results in death
Is life-threatening (defined as an event in which the participant was at risk of death at the time of the event; it does not refer to an event which hypothetically might have caused death if it were more severe)
Requires inpatient hospitalization or causes prolongation of existing hospitalization (see NOTE below) NOTE: The following hospitalizations are not considered SAEs in BMS clinical studies: o a visit to the emergency room or other hospital department < 24 hours, that does not result in admission (unless considered an important medical or life-threatening event) o elective surgery, planned prior to signing consent o admissions as per protocol for a planned medical/surgical procedure o routine health assessment requiring admission for baseline/trending of health status (e.g., routine colonoscopy) o medical/surgical admission other than to remedy ill health and planned prior to entry into the study. Appropriate documentation is required in these cases o admission encountered for another life circumstance that carries no bearing on health status and requires no medical/surgical intervention (e.g., lack of housing, economic inadequacy, caregiver respite, family circumstances, administrative reason) o admission for administration of anticancer therapy in the absence of any other SAEs (applies to oncology protocols)
Results in persistent or significant disability/incapacity
Is a congenital anomaly/birth defect
is an important medical event (defined as a medical event(s) that may not be immediately life-threatening or result in death or hospitalization but, based upon appropriate medical and scientific judgment, may jeopardize the participant or may require intervention [e.g., medical, surgical] to prevent one of the other serious outcomes listed in the definition above.) Examples of such events include, but are not limited to, intensive treatment in an emergency room or at home for allergic bronchospasm; blood dyscrasias or convulsions that do not result in hospitalization.) Potential drug induced liver injury (DILI) is also considered an important medical event. (See Section 9.2.7 for the definition of potential DILI.)

Suspected transmission of an infectious agent (e.g., pathogenic or nonpathogenic) via the study treatment is an SAE.

Although pregnancy, overdose, cancer, and potential drug induced liver injury (DILI) are not always serious by regulatory definition, these events must be handled as SAEs. (See Section 9.2.5 for reporting pregnancies).

Any component of a study endpoint that is considered related to study therapy should be reported as SAE (e.g., death is an endpoint, if death occurred due to anaphylaxis, anaphylaxis must be reported).

EVALUATING AES AND SAES

Assessment of Causality
The causal relationship to study drug is determined by a physician and should be used to assess all adverse events (AE). The causal relationship can be one of the following: Related: There is a reasonable causal relationship between study drug administration and the AE. Not related: There is not a reasonable causal relationship between study drug administration and the AE. The term "reasonable causal relationship" means there is evidence to suggest a causal relationship.

Follow-up of AEs and SAEs
If only limited information is initially available, follow-up reports are required. (Note: Follow-up SAE reports must include the same investigator term(s) initially reported.) If an ongoing SAE changes in its intensity or relationship to study drug or if new information becomes available, the SAE report must be updated and submitted within 24 hours to BMS (or designee) using the same procedure used for transmitting the initial SAE report. All SAEs must be followed to resolution or stabilization.

REPORTING OF SAES TO SPONSOR OR DESIGNEE

- SAEs, whether related or not related to study drug, and pregnancies must be reported to BMS (or designee) within 24 hours of awareness of the event.
- SAEs must be recorded on the SAE Report Form; pregnancies on a Pregnancy Surveillance Form (electronic or paper forms).
- The preferred method for SAE data reporting collection is through the eCRF.
- The paper SAE/pregnancy surveillance forms are only intended as a back-up option when the eCRF system is not functioning.
    - In this case, the paper forms are to be transmitted via email or confirmed facsimile (fax) transmission to:

SAE Email Address: Refer to Contact Information list.

SAE Facsimile Number: Refer to Contact Information list.

For studies capturing SAEs through electronic data capture (EDC), electronic submission is the required method for reporting. In the event the electronic system is unavailable for transmission, paper forms must be used and submitted immediately. When paper forms are used, the original paper forms are to remain on site.

SAE Telephone Contact (required for SAE and pregnancy reporting): Refer to Contact Information list

APPENDIX 4 WOMEN OF CHILDBEARING POTENTIAL DEFINITIONS AND METHODS OF CONTRACEPTION

DEFINITIONS

Woman of Childbearing Potential (WOCBP)

A woman is considered fertile following menarche and until becoming post-menopausal unless permanently sterile. Permanent sterilization methods include hysterectomy, bilateral salpingectomy, and bilateral oophorectomy.

Women in the following categories are not considered WOCBP

- Premenarchal
- Premenopausal female with 1 of the following:
  - Documented hysterectomy
  - Documented bilateral salpingectomy
  - Documented bilateral oophorectomy

Note: Documentation can come from the site personnel's review of the participant's medical records, medical examination, or medical history interview.

- Postmenopausal female
  - A postmenopausal state is defined as 12 months of amenorrhea in a woman over age 45 years in the absence of other biological or physiological causes. In addition, females under the age of 55 years must have a serum follicle stimulating hormone, (FSH) level > 40 mIU/mL to confirm menopause.

CONTRACEPTION GUIDANCE FOR FEMALE PARTICIPANTS OF CHILD BEARING POTENTIAL

One of the highly effective methods of contraception listed below is required during study duration and until the end of relevant systemic exposure, defined as 5 months after the end of study treatment.

Local laws and regulations may require use of alternative and/or additional contraception methods.

Highly Effective Contraceptive Methods That Are User Dependent *Failure rate of <1% per year when used consistently and correctly.*[a]
• Combined (estrogen- and progestogen-containing) hormonal contraception associated with inhibition of ovulation[b] – oral – intravaginal – transdermal

• Progestogen-only hormonal contraception associated with inhibition of ovulation[b] – oral – injectable
Highly Effective Methods That Are User Independent
• Implantable progestogen-only hormonal contraception associated with inhibition of ovulation [b] • Hormonal methods of contraception including oral contraceptive pills containing a combination of estrogen and progesterone, vaginal ring, injectables, implants and intrauterine hormone-releasing system (IUS)[c] • Intrauterine device (IUD)[c] • Bilateral tubal occlusion
• Vasectomized partner *A vasectomized partner is a highly effective contraception method provided that the partner is the sole male sexual partner of the WOCBP and the absence of sperm has been confirmed. If not, an additional highly effective method of contraception should be used.*
• Sexual abstinence *Sexual abstinence is considered a highly effective method only if defined as refraining from heterosexual intercourse during the entire period of risk associated with the study drug. The reliability of sexual abstinence needs to be evaluated in relation to the duration of the study and the preferred and usual lifestyle of the participant.* • It is not necessary to use any other method of contraception when complete abstinence is elected. • WOCBP participants who choose complete abstinence must continue to have pregnancy tests, as specified in Section 2. • Acceptable alternate methods of highly effective contraception must be discussed in the event that the WOCBP participants chooses to forego complete abstinence
NOTES: [a] Typical use failure rates may differ from those when used consistently and correctly. Use should be consistent with local regulations regarding the use of contraceptive methods for participants participating in clinical studies. [b] Hormonal contraception may be susceptible to interaction with the study drug, which may reduce the efficacy of the contraceptive method. Hormonal contraception is permissible only when there is sufficient evidence that the IMP and other study medications will not alter hormonal exposures such that contraception would be ineffective or result in increased exposures that could be potentially hazardous. In this case, alternative methods of contraception should be utilized. [c] Intrauterine devices and intrauterine hormone releasing systems are acceptable methods of contraception in the absence of definitive drug interaction studies when hormone exposures from intrauterine devices do not alter contraception effectiveness

Unacceptable Methods of Contraception
• Male or female condom with or without spermicide. Male and female condoms cannot be used simultaneously

- Diaphragm with spermicide
- Cervical cap with spermicide
- Vaginal Sponge with spermicide
- Progestogen-only oral hormonal contraception, where inhibition of ovulation is not the primary mechanism of action
- Periodic abstinence (calendar, symptothermal, post-ovulation methods)
- Withdrawal (coitus interruptus).
- Spermicide only
- Lactation amenorrhea method (LAM)

CONTRACEPTION GUIDANCE FOR MALE PARTICIPANTS WITH PARTNER(S) OF CHILD BEARING POTENTIAL.

Male participants with female partners of childbearing potential are eligible to participate if they agree to the following during the treatment and until the end of relevant systemic exposure.

- Inform any and all partner(s) of their participation in a clinical drug study and the need to comply with contraception instructions as directed by the investigator.
- Male participants are required to use a condom for study duration and until end of relevant systemic exposure defined as 7 months after the end of study treatment.
- Female partners of males participating in the study to consider use of effective methods of contraception until the end of relevant systemic exposure, defined as 7 months after the end of treatment in the male participant.
- Male participants with a pregnant or breastfeeding partner must agree to remain abstinent from penile vaginal intercourse or use a male condom during each episode of penile penetration during the treatment and until 7 months after the end of study treatment.
- Refrain from donating sperm for the duration of the study treatment and until 7 months after the end of study treatment.

COLLECTION OF PREGNANCY INFORMATION

Guidance for collection of Pregnancy Information and outcome of pregnancy on the Pregnancy Surveillance Form is provided in Section 9.2.5 and the Appendix for Adverse Events and Serious Adverse Events Definitions and procedures for Evaluating, Follow-up and Reporting

APPENDIX 5 MANAGEMENT ALGORITHMS

These general guidelines constitute guidance to the Investigator and may be supplemented by discussions with the Medical Monitor representing the Sponsor. The guidance applies to all immuno-oncology agents and regimens.

A general principle is that differential diagnoses should be diligently evaluated according to standard medical practice. Non-inflammatory etiologies should be considered and appropriately treated.

Corticosteroids are a primary therapy for immuno-oncology drug-related adverse events. The oral equivalent of the recommended IV doses may be considered for ambulatory patients with low-grade toxicity. The lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.

Consultation with a medical or surgical specialist, especially prior to an invasive diagnostic or therapeutic procedure, is recommended.

The frequency and severity of the related adverse events covered by these algorithms will depend on the immuno-oncology agent or regimen being used.

GI Adverse Event Management Algorithm

Rule out non-inflammatory causes. If non-inflammatory cause is identified, treat accordingly and continue I-O therapy. Opiates/narcotics may mask symptoms of perforation. Infliximab should not be used in cases of perforation or sepsis.

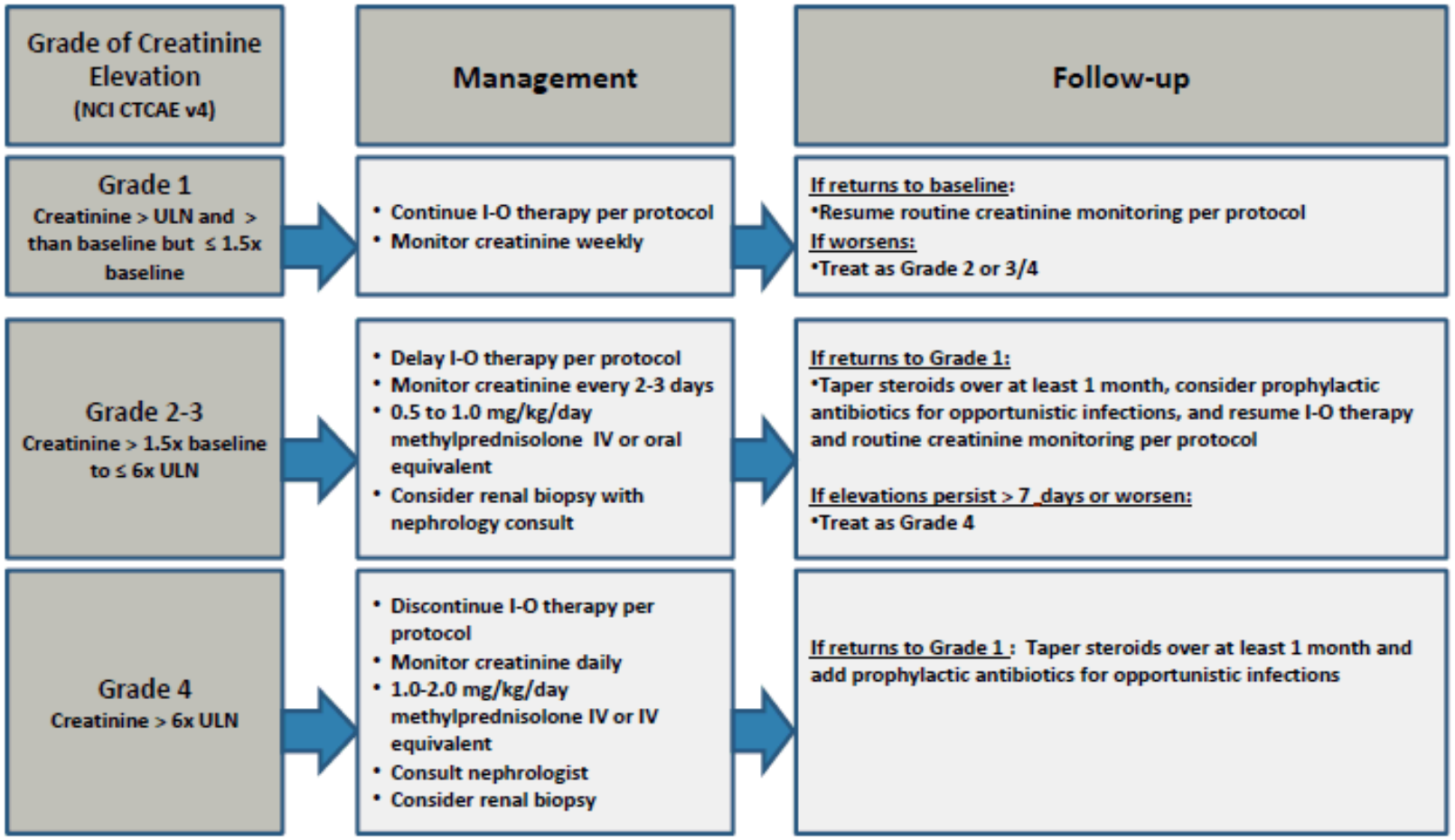
Grade of Diarrhea/ Colitis (NCI CTCAE v4)	Management	Follow-up
Grade 1 Diarrhea:< 4 stools/day over baseline; Colitis: asymptomatic	• Continue I-O therapy per protocol • Symptomatic treatment	• Close monitoring for worsening symptoms. • Educate patient to report worsening immediately If worsens: • Treat as Grade 2 or 3/4
Grade 2 Diarrhea: 4-6 stools per day over baseline; IV fluids indicated <24 hrs; not interfering with ADL Colitis: abdominal pain; blood in stool	• Delay I-O therapy per protocol • Symptomatic treatment	If improves to grade 1: • Resume I-O therapy per protocol If persists > 5-7 days or recur: • 0.5-1.0 mg/kg/day methylprednisolone or oral equivalent • When symptoms improve to grade 1, taper steroids over at least 1 month, consider prophylactic antibiotics for opportunistic infections, and resume I-O therapy per protocol. If worsens or persists > 3-5 days with oral steroids: • Treat as grade 3/4
Grade 3-4 Diarrhea (G3): ≥7 stools per day over baseline; incontinence; IV fluids ≥24 hrs; interfering with ADL Colitis (G3): severe abdominal pain, medical intervention indicated, peritoneal signs G4: life-threatening, perforation	• Discontinue I-O therapy per protocol • 1.0 to 2.0 mg/kg/day methylprednisolone IV or IV equivalent • Add prophylactic antibiotics for opportunistic infections • Consider lower endoscopy	If improves: • Continue steroids until grade 1, then taper over at least 1 month If persists > 3-5 days, or recurs after improvement: • Add infliximab 5 mg/kg (if no contraindication). Note: Infliximab should not be used in cases of perforation or sepsis

Patients on IV steroids may be switched to an equivalent dose of oral corticosteroids (e.g. prednisone) at start of tapering or earlier, once sustained clinical improvement is observed. Lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.

Updated 05-Jul-2016

Renal Adverse Event Management Algorithm

Rule out non-inflammatory causes. If non-inflammatory cause, treat accordingly and continue I-O therapy

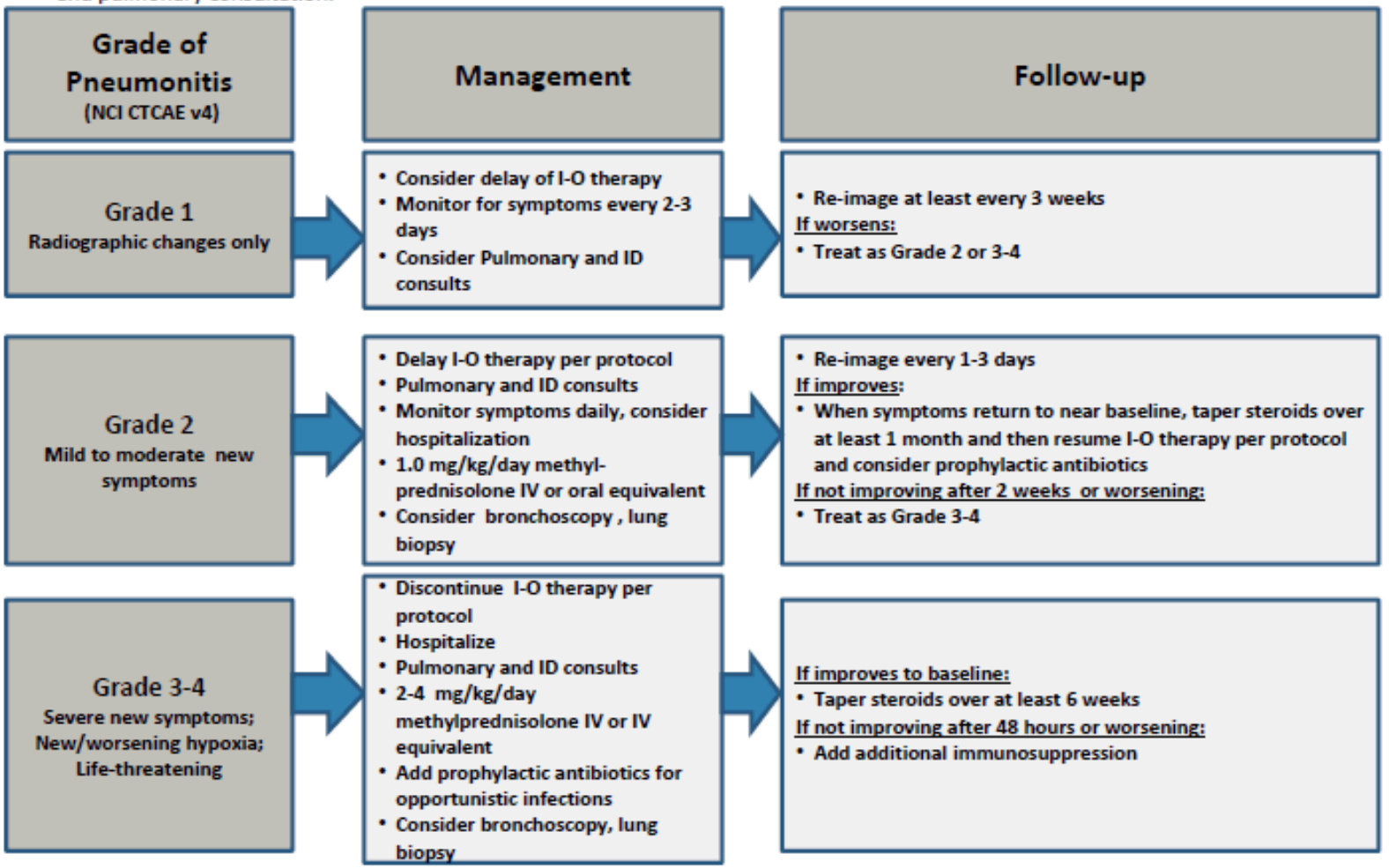


Patients on IV steroids may be switched to an equivalent dose of oral corticosteroids (e.g. prednisone) at start of tapering or earlier, once sustained clinical improvement is observed. Lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.

Updated 05-Jul-2016

Pulmonary Adverse Event Management Algorithm

Rule out non-inflammatory causes. If non-inflammatory cause, treat accordingly and continue I-O therapy. Evaluate with imaging and pulmonary consultation.

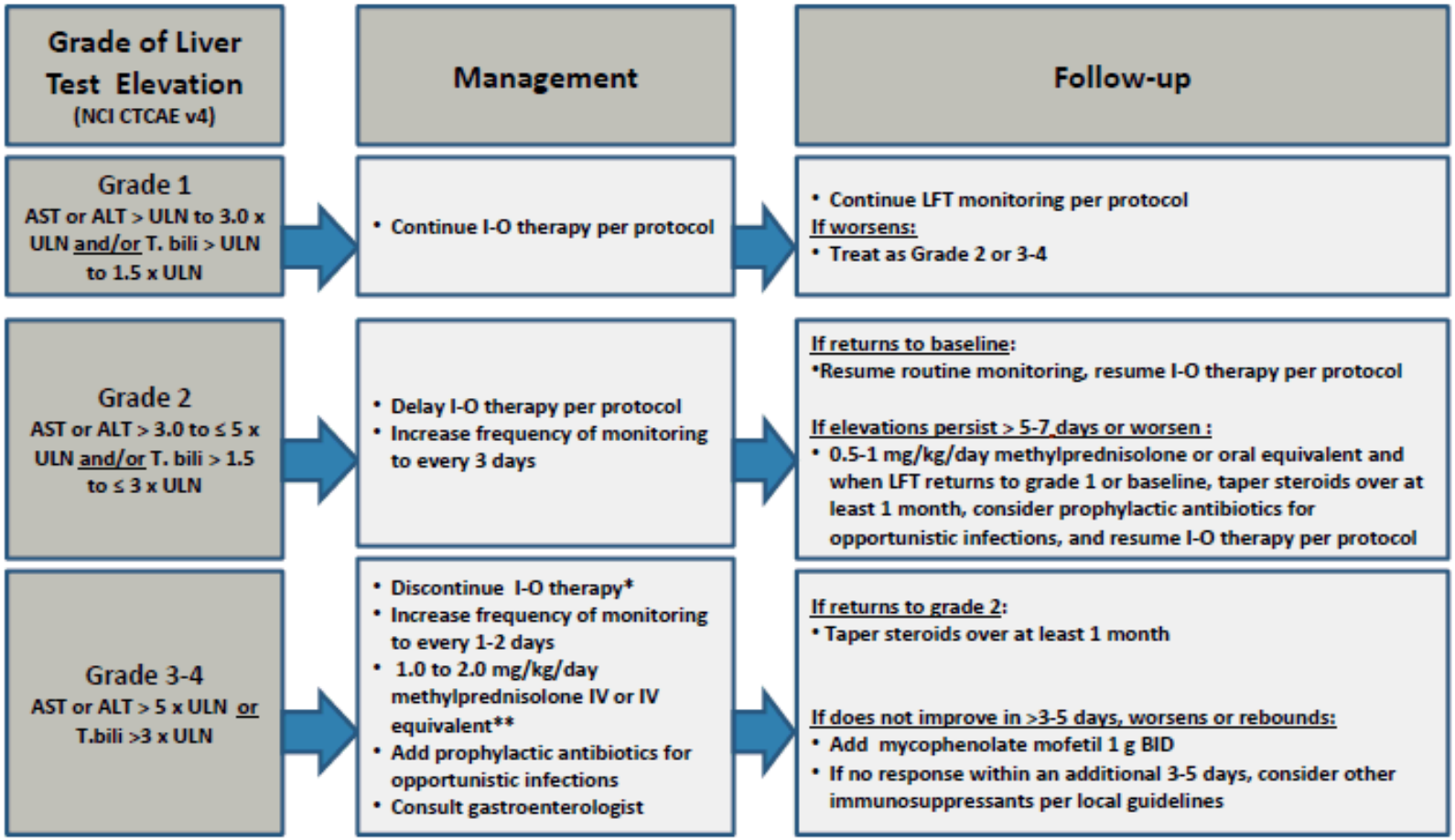


Patients on IV steroids may be switched to an equivalent dose of oral corticosteroids (e.g. prednisone) at start of tapering or earlier, once sustained clinical improvement is observed. Lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.

Updated 05-Jul-2016

Hepatic Adverse Event Management Algorithm

Rule out non-inflammatory causes. If non-inflammatory cause, treat accordingly and continue I-O therapy. Consider imaging for obstruction.



Patients on IV steroids may be switched to an equivalent dose of oral corticosteroids (e.g. prednisone) at start of tapering or earlier, once sustained clinical improvement is observed. Lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.

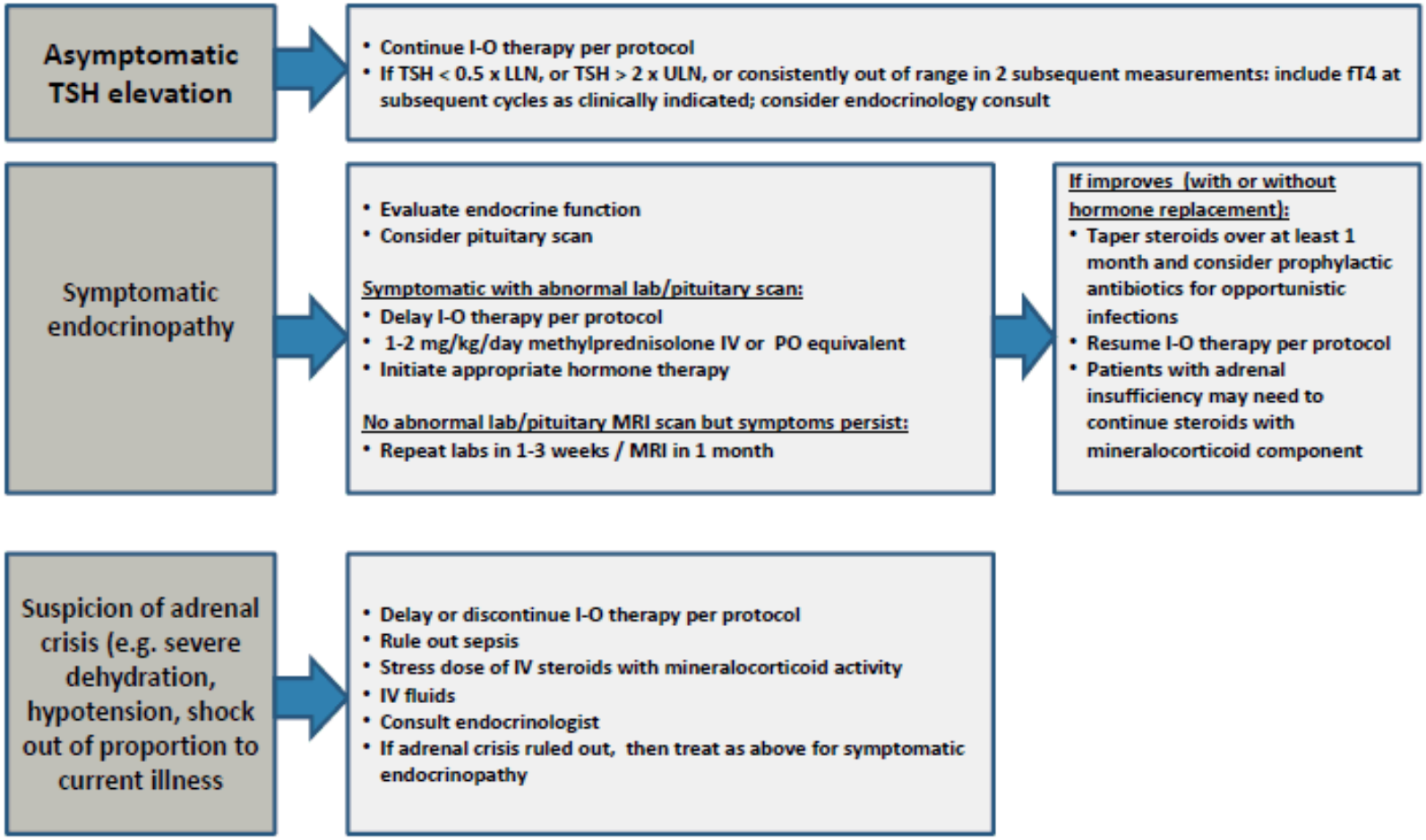
***I-O therapy may be delayed rather than discontinued if AST/ALT ≤ 8 x ULN or T.bili ≤ 5 x ULN.**

****The recommended starting dose for grade 4 hepatitis is 2 mg/kg/day methylprednisolone IV.**

Updated 05-Jul-2016

Endocrinopathy Management Algorithm

Rule out non-inflammatory causes. If non-inflammatory cause, treat accordingly and continue I-O therapy. Consider visual field testing, endocrinology consultation, and imaging.

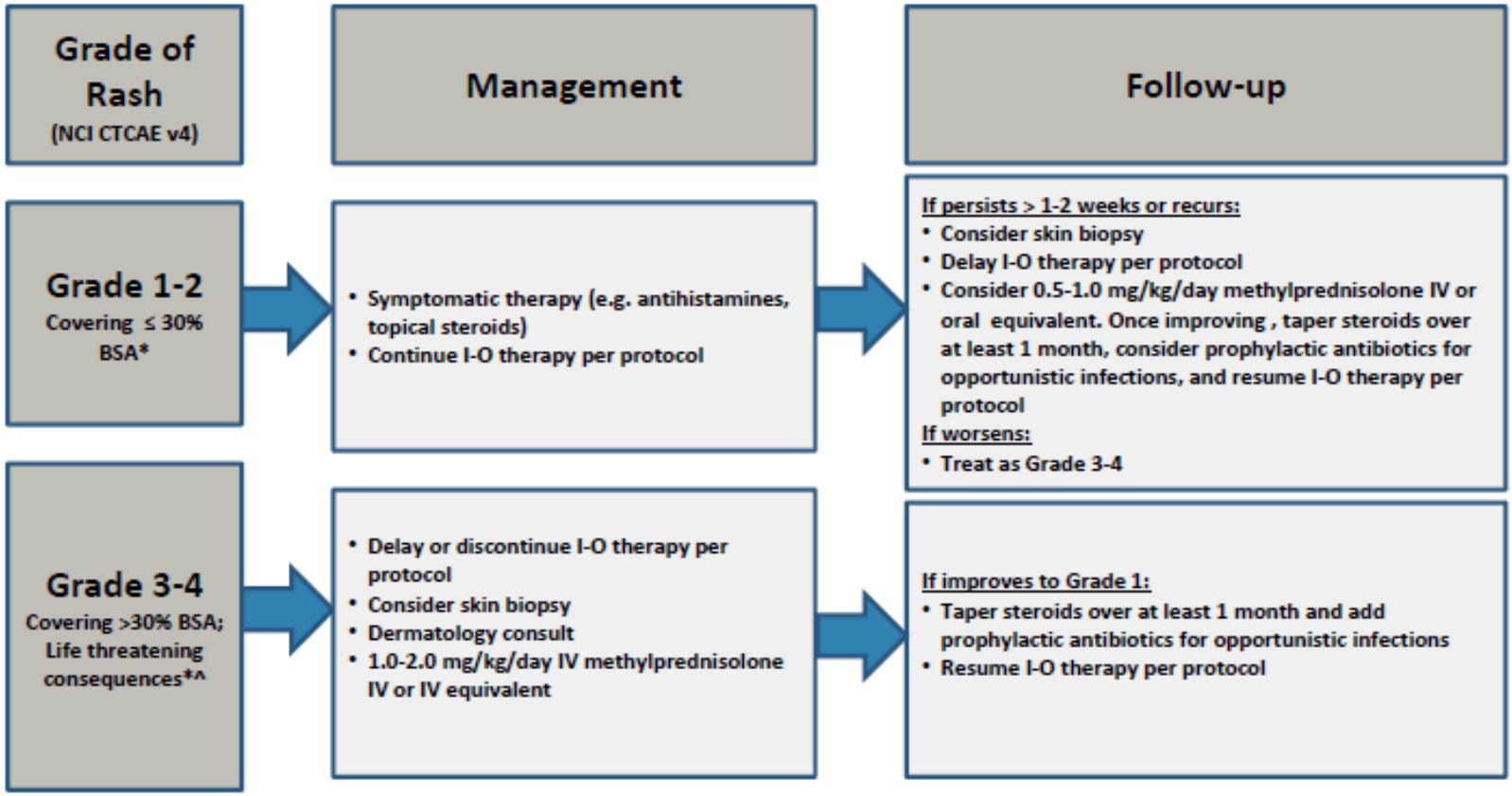


Patients on IV steroids may be switched to an equivalent dose of oral corticosteroids (e.g. prednisone) at start of tapering or earlier, once sustained clinical improvement is observed. Lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.

Updated 05-Jul-2016

Skin Adverse Event Management Algorithm

Rule out non-inflammatory causes. If non-inflammatory cause, treat accordingly and continue I-O therapy.

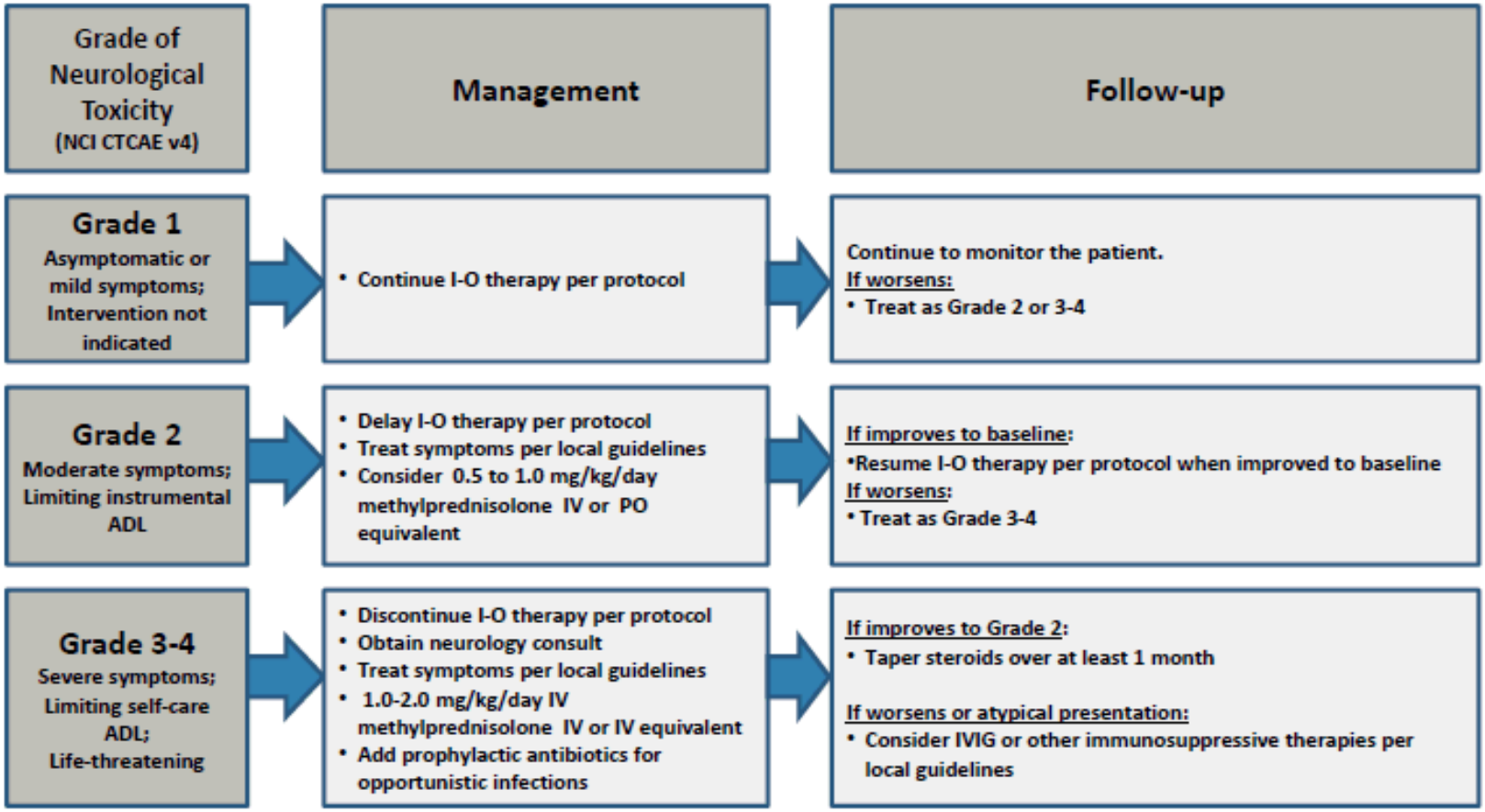


Patients on IV steroids may be switched to an equivalent dose of oral corticosteroids (e.g. prednisone) at start of tapering or earlier, once sustained clinical improvement is observed. Lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.
*Refer to NCI CTCAE v4 for term-specific grading criteria.
^If SJS/TEN is suspected, withhold I-O therapy and refer patient for specialized care for assessment and treatment. If SJS or TEN is diagnosed, permanently discontinue I-O therapy.

Updated 05-Jul-2016

Neurological Adverse Event Management Algorithm

Rule out non-inflammatory causes. If non-inflammatory cause, treat accordingly and continue I-O therapy.



Patients on IV steroids may be switched to an equivalent dose of oral corticosteroids (e.g. prednisone) at start of tapering or earlier, once sustained clinical improvement is observed. Lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.

Updated 05-Jul-2016

APPENDIX 6 INTERNATIONAL METASTATIC RCC DATABASE CONSORTIUM (IMDC) PROGNOSTIC CRITERIA

Adverse Prognostic Factors
Clinical
KPS < 80% Time from initial diagnosis (including original localized disease if applicable) to treatment < 1 year
Laboratory
Hemoglobin < LLN Corrected calcium > 10 mg/dL Absolute neutrophil count > ULN Platelet count > ULN

Note: The corrected calcium criterion was adapted from to Heng et al, 2009 to account for local laboratories that may not provide an ULN for corrected calcium.

Abbreviations: KPS= Karnofsky Performance Status; LLN = Lower limit of normal; ULN = Upper limit of normal

Corrected calcium (mg/dL) = measured total Ca (mg/dL) + 0.8 (4.0 - serum albumin [g/dL]), where 4.0 represents the average albumin level in g/dL.

Corrected calcium (mmol/L) = measured total Ca (mmol/L) + 0.02 (40 - serum albumin [g/L]), where 40 represents the average albumin level in g/L

Risk Group Based on Number of Adverse Prognostic Factors	
Number of Adverse Prognostic Factors Present	**Risk Group**
0	Favorable
1-2	Intermediate
3-6	Poor

Reference: Heng D, Xie W, Regan M, et al. Prognostic factors for overall survival in patients with metastatic renal cell carcinoma treated with vascular endothelial growth factor-targeted agents: results from a large, multicenter study. J Clin Oncol 2009; 27(34):5794-5799.

APPENDIX 7 PERFORMANCE STATUS SCORES

STATUS	SCALES		STATUS
	KARNOFSKY	ZUBROD-ECOG-WHO	
Normal, no complaints	100	0	Normal activity
Able to carry on normal activities Minor signs or symptoms of disease	90	0	Symptoms, but fully ambulatory
Normal activity with effort	80	1	
Cares for self. Unable to carry on normal activity or to do active work	70	1	Symptomatic, but in bed < 50% of the day.
Requires occasional assistance, but able to care for most of his needs	60	2	
Requires considerable assistance and frequent medical care	50	2	Needs to be in bed > 50% of the day, but not bedridden
Disabled. Requires special care and assistance	40	3	
Severely disabled. Hospitalization indicated though death non imminent	30	3	Unable to get out of bed
Very sick. Hospitalization necessary. Active supportive treatment necessary	20	4	
Moribund	10	4	
Dead	0	5	Dead

APPENDIX 8 RADIOLOGIC EVALUATION CRITERIA IN SOLID TUMOURS VERSION 1.1 (RECIST CRITERIA 1.1)

1 ASSESSMENT OF OVERALL TUMOR BURDEN AND MEASURABLE DISEASE

Subjects must have measureable disease to be eligible for this study.

To assess objective response or future progression, it is necessary to estimate the *overall tumor burden at baseline* and use this as a comparator for subsequent measurements. Measurable disease is defined by the presence of at least 1 measurable tumor lesion. When computed tomography (CT) scans have slice thickness greater than 5 mm, the minimum size for a measurable lesion should be twice the slice thickness.

At baseline, tumor lesions/lymph nodes will be categorized measurable or nonmeasurable, which are discussed below.

1.1 Measurable Lesions

Measurable lesions must be accurately measured in at least 1 dimension (longest diameter in the plane of the measurement to be recorded) with a minimum size of:

- 10 mm by CT/magnetic resonance imaging (MRI) scan - (CT/MRI scan slice thickness no greater than 5 mm)
- 10 mm caliper measurement by clinical exam (lesions which cannot be accurately measured with calipers should be recorded as nonmeasurable)
- 20 mm by chest x-ray
- Malignant lymph nodes: To be considered pathologically enlarged and measurable, a lymph node must be ≥ 15 mm in short axis when assessed by CT scan (CT scan slice thickness recommended to be no greater than 5 mm). At baseline and in follow-up, only the short axis will be measured and followed.

1.2 Non-measurable Lesions

- All other lesions, including small lesions (longest diameter < 10 mm or pathological lymph nodes with ≥ 10 to < 15 mm short axis), as well as truly non-measurable lesions
- Lesions considered truly nonmeasurable include: leptomeningeal disease, ascites, pleural or pericardial effusion, inflammatory breast disease, lymphangitic involvement of skin or lung, and abdominal masses/abdominal organomegaly identified by physical exam that is not measurable by reproducible imaging techniques.

1.3 Special Considerations Regarding Lesion Measurability

1.3.1 Bone Lesions

- Bone scan, positron emission tomography (PET) scan, or plain films are not considered adequate imaging techniques to measure bone lesions. However, these techniques can be used to confirm the presence or disappearance of bone lesions.
- Lytic bone lesions or mixed lytic-blastic lesions, with identifiable soft tissue components that can be evaluated by cross sectional imaging techniques such as CT or MRI, can be considered

measurable lesions if the soft tissue component meets the definition of measurability described above.

- Blastic bone lesions are nonmeasurable.

1.3.2 Cystic Lesions

- Lesions that meet the criteria for radiographically defined simple cysts should not be considered as malignant lesions (neither measurable nor nonmeasurable) since they are, by definition, simple cysts.
- 'Cystic lesions' thought to represent cystic metastases can be considered as measurable lesions if they meet the definition of measurability described above. However, if non-cystic lesions are present in the same subject, these are preferred for selection as target lesions.

1.3.3 Lesions with Prior Local Treatment

Tumor lesions situated in a previously irradiated area or in an area subjected to other locoregional therapy are usually not considered measurable, unless there has been demonstrated progression in the lesion. Measurable lesions may be in an irradiated field as long as there is documented progression, and the lesion(s) can be reproducibly measured.

1.4 Specifications by Methods of Measurements

1.4.1 Measurement of Lesions

All measurements should be recorded in metric notation (mm). All baseline evaluations should be performed as close as possible to the treatment start and never more than 30 days before the beginning of the treatment.

1.4.2 Method of Assessment

The **same method of assessment and the same technique should be used** to characterize each identified and reported lesion at baseline and during follow-up. Imaging-based evaluation should always be done rather than clinical examination, unless the lesion(s) being followed cannot be imaged but are assessable by clinical exam.

1.4.2.1 CT/MRI Scan

CT/MRI is the best currently available and reproducible method to measure lesions selected for response assessment. Measurability of lesions on CT/MRI scan is based on the assumption that CT/MRI slice thickness is 5 mm or less. When CT scans have slice thickness greater than 5 mm, the minimum size for a measurable lesion should be twice the slice thickness.

1.4.2.2 Chest X-ray

Chest CT is preferred over chest X-ray, particularly when progression is an important endpoint, since CT is more sensitive than X-ray in identifying new lesions. However, lesions on chest X-ray may be considered measurable if they are clearly defined and surrounded by aerated lung.

1.4.2.3 Clinical Lesions

Clinical lesions will only be considered measurable when they are superficial and ≥ 10 mm diameter as assessed using calipers. For the case of skin lesions, documentation by color

photography including a ruler to estimate the size of the lesion is suggested. As previously noted, when lesions can be evaluated by both clinical examination and imaging, imaging evaluation should be undertaken since it is more objective and may also be reviewed at the end of the study.

1.4.2.4 Ultrasound

Ultrasound is **not** useful in assessment of lesion size and should not be used as a method of measurement. If new lesions are identified by ultrasound in the course of the study, confirmation by CT or MRI is advised.

1.4.2.5 Endoscopy, Laparoscopy

The utilization of these techniques for objective tumor evaluation is **not** advised.

1.4.2.6 Tumor Markers

Tumor markers **alone** cannot be used to assess objective tumor response.

2 BASELINE DOCUMENTATION OF 'TARGET' AND 'NON-TARGET' LESIONS

2.1 Target Lesions

When more than 1 measurable lesion is present at baseline, all lesions up to **a maximum of 5 lesions total (and a maximum of 2 lesions per organ) representative of all involved organs should be identified as *target lesions*** and will be recorded and measured at baseline.

Target lesions should be selected on the basis of their **size** (lesions with the longest diameter), be representative of all involved organs, and should lend themselves to **reproducible repeated measurements.**

A **sum of the diameters** (longest for non-nodal lesions, short axis for nodal lesions) for all target lesions will be calculated and reported as the **baseline sum diameters.** If lymph nodes are to be included in the sum, then as noted below, only the **short** axis is added into the sum. The baseline sum diameters will be used as reference to further characterize any objective tumor regression in the measurable dimension of the disease.

2.1.1 Lymph Nodes

Lymph nodes merit special mention since they are normal anatomical structures which may be visible by imaging even if not involved by tumor. Pathological nodes, which are defined as measurable and may be identified as target lesions, must meet the criterion of a **short axis of ≥ 15 mm by CT scan**. Only the short axis of these nodes will contribute to the baseline sum. Nodes that have a short axis < 10 mm are considered nonpathological and should not be recorded or followed.

2.2 Non-target Lesions

All other lesions (or sites of disease), including pathological lymph nodes, should be identified as **non-target lesions** and should also be recorded at baseline. Measurements are not required, and these lesions should be followed as **'present', 'absent', or in rare cases, 'unequivocal progression'.** In addition, it is possible to record multiple non-target lesions involving the same

organ as a single item on the case record form (e.g., 'multiple enlarged pelvic lymph nodes' or 'multiple liver metastases').

3 TUMOR RESPONSE EVALUATION

3.1 Evaluation of Target Lesions

- Complete Response (CR): **Disappearance of all target lesions.** Any pathological lymph nodes (whether target or nontarget) must have reduction in short axis to < 10 mm.
- Partial Response (PR): At least a **30% decrease in the sum of diameters of target lesions,** taking as reference the baseline sum diameters.
- Progressive Disease (PD): At least a **20% increase in the sum of diameters of target lesions,** taking as reference the **smallest sum on study** (this includes the baseline sum if that is the smallest on study). In addition to the relative increase of 20%, the sum must also demonstrate an **absolute increase of at least 5 mm**. (Note*:* The appearance of 1 or more new lesions is also considered progression).
- Stable Disease (SD): Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD, taking as reference the smallest sum on study.

3.1.1 Special Notes on the Assessment of Target Lesions

3.1.1.1 Lymph Nodes

Lymph nodes merit special mention since they are normal anatomical structures which may be visible by imaging even if not involved by tumor. Pathological nodes, which are defined as measurable and may be identified as target lesions, must meet the criterion of a **short axis of ≥ 15 mm by CT scan**. Only the short axis of these nodes will contribute to the baseline sum. Nodes that have a short axis <10 mm are considered nonpathological and should not be recorded or followed.

3.1.1.2 Target Lesions That Become 'Too Small to Measure'

All lesions (nodal and nonnodal) recorded at baseline should have their actual measurements recorded at each subsequent evaluation, even when very small (e.g., 2 mm). If the radiologist is able to provide an actual measurement, that should be recorded even if it is below 5 mm.

However, when such a lesion becomes difficult to assign an exact measure to then:

- If it is the opinion of the radiologist that the lesion has likely disappeared, the measurement should be recorded as 0 mm.
- If the lesion is believed to be present and is faintly seen but too small to measure, a default value of 5 mm should be assigned (Note: In case of a lymph node believed to be present and faintly seen but too small to measure, a default value of 5 mm should be assigned in this circumstance as well). This default value is derived from the 5 mm CT slice thickness (but should not be changed with varying CT slice thickness).

3.1.1.3 Target Lesions that Split or Coalesce on Treatment

- When non-nodal lesions fragment, the longest diameters of the fragmented portions should be added together to calculate the target lesion sum.

- As lesions coalesce, a plane between them may be maintained that would aid in obtaining maximal diameter measurements of each individual lesion. If the lesions have truly coalesced such that they are no longer separable, the vector of the longest diameter in this instance should be the maximal longest diameter for the 'coalesced lesion'

3.2 Evaluation of Non-target Lesions

While some non-target lesions may actually be measurable, they need not be measured and instead should be assessed only qualitatively at the time points specified in the protocol.

- CR: Disappearance of all non-target lesions. All lymph nodes must be nonpathological in size (< 10 mm short axis).
- PD: Unequivocal progression of existing non-target lesions. (Note: The appearance of 1 or more new lesions is also considered progression).
- NonCR/NonPD: Persistence of 1 or more non-target lesion(s).

3.2.1 Special Notes on Assessment of Non-target Lesions

The concept of progression of non-target disease requires additional explanation as discussed below.

3.2.1.1 When the Subject Also has Measurable Disease

- To achieve unequivocal progression on the basis of the non-target disease, there must be an overall level of substantial worsening in non-target disease, such that even in presence of SD or PR in target disease, the overall tumor burden has increased sufficiently to merit discontinuation of therapy.
- A modest 'increase' in the size of 1 or more non-target lesions is usually not sufficient to quality for unequivocal progression status.

3.2.1.2 When the Subject has Only Non-measurable Disease

- To achieve unequivocal progression on the basis of the non-target disease, there must be an overall level of substantial worsening, such that the overall tumor burden has increased sufficiently to merit discontinuation of therapy.
- A modest increase in the size of 1 or more non-target lesions is usually not sufficient to qualify for unequivocal progression status.
- Because worsening in non-target disease cannot be easily quantified (by definition, if all lesions are nonmeasurable), a useful test that can be applied when assessing subjects for unequivocal progression is to consider if the increase in overall disease burden based on the change in non-measurable disease is comparable in magnitude to the increase that would be required to declare PD for measurable disease, ie, an increase in tumor burden representing an additional 73% increase in volume (which is equivalent to a 20% increase diameter in a measurable lesion). Examples include an increase in a pleural effusion from 'trace' to 'large', an increase in lymphangitic disease from localized to widespread, or may be described in protocols as 'sufficient to require a change in therapy'.
- If unequivocal progression is seen, the subject should be considered to have had overall PD at that point.

3.2.1.3 Tumor Markers

Tumor markers will not be used to assess objective tumor responses.

3.3 New Lesions

The appearance of new malignant lesions denotes disease progression. The finding of a new lesion should be unequivocal, that is, not attributable to differences in scanning technique, change in imaging modality, or findings thought to represent something other than tumor (for example, some new bone lesions may be simply healing or flare of preexisting lesions). This is particularly important when the subject's baseline lesions show PR or CR. For example, necrosis of a liver lesion may be reported on a CT scan report as a new cystic lesion, which it is not.

A lesion identified on a follow-up study in an anatomical location that was **not** scanned at baseline is considered a new lesion and will indicate disease progression. An example of this is the subject who has visceral disease at baseline and while on study, has a CT or MRI brain scan ordered that reveals metastases. The subject's brain metastases are considered to be evidence of PD even if he/she did not have brain imaging at baseline.

If a new lesion is equivocal, for example because of its small size, continued therapy and follow-up evaluation will clarify if it represents truly new disease. If repeat scans confirm there is definitely a new lesion, then progression should be declared using the date of the initial scan.

4 RESPONSE CRITERIA

4.1 Time Point Response

A response assessment should occur at each time point specified in the protocol.

For subjects who have **measurable disease** at baseline, Table 1 provides a summary of the overall response status calculation at each time point.

Table 1: Subjects with Target (+/– Non-target) Disease

Target Lesions	Non-target Lesions	New Lesions	Overall Response
CR	CR	No	CR
CR	NonCR/nonPD	No	PR
CR	Not evaluated	No	PR
PR	NonPD or not all evaluated	No	PR
SD	NonPD or not all evaluated	No	SD
Not all evaluated	NonPD	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD

Abbreviations: CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, NE = not evaluable.

4.1.1 Missing Assessments and Not Evaluable Designation

When no imaging/measurement is done at a particular time point, the subject is **not evaluable (NE)** at that time point. If only a subset of lesion measurements are made at an assessment, the case is also considered NE at that time point, unless a convincing argument can be made that the contribution of the individual missing lesion(s) would not have changed the assigned time point response.

4.2 Best Overall Response: All Time Points

The best overall response is determined once all the data for the subject is known. It is the best response recorded from the start of the study treatment to the date of radiographic progression per RECIST 1.1 or the date of subsequent anti-cancer therapy (including tumor-directed radiotherapy and tumor-directed surgery), whichever occurs first, taking into account any requirement for confirmation. The subject's best overall response assignment will depend on the findings of both target and non-target disease and will also take into consideration the appearance of new lesions.

Complete or partial responses may be claimed only if the criteria for each are met at a subsequent time point of ≥ 4 weeks later. In this circumstance, the best overall response can be interpreted as specified in Table 2.

When SD is believed to be best response, it must meet the protocol specified minimum time from baseline. In this protocol, measurements must have met the SD criteria at least once after study entry at a minimum interval of 6 weeks from randomization in order for SD to be the best response.

Table 2: Best Overall Response When Confirmation of CR and PR is Required

Overall Response	Overall Response	Best Overall Response
First Time Point	Subsequent Time Point	
CR	CR	CR
CR	PR	SD, PD, or PR[a]
CR	SD	SD provided minimum criteria for SD duration met, otherwise, PD
CR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
CR	NE	SD provided minimum criteria for SD duration met, otherwise NE
PR	CR	PR
PR	PR	PR
PR	SD	SD
PR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
PR	NE	SD provided minimum criteria for SD duration met, otherwise NE

Table 2: Best Overall Response When Confirmation of CR and PR is Required

Overall Response	Overall Response	Best Overall Response
NE	NE	NE

[a] Abbreviations: CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, NE = not evaluable.

If a CR is truly met at first time point, then any disease seen at a subsequent time point, even disease meeting PR criteria relative to baseline, makes the disease PD at that point (since disease must have reappeared after CR). Best response would depend on whether minimum duration for SD was met. However, sometimes 'CR' may be claimed when subsequent scans suggest small lesions were likely still present and in fact the subject had PR, not CR at the first time point. Under these circumstances, the original CR should be changed to PR and the best response is PR.

4.2.1 Confirmation Scans

Verification of Response: To be assigned a status of CR or PR, changes in tumor measurements must be confirmed by consecutive repeat assessments that should be performed no less than 28 days after the criteria for response are first met. For this study, the next scheduled tumor assessment can meet this requirement.

4.3 Duration of Response

4.3.1 Duration of Overall Response

The duration of overall response is measured from the time measurement criteria are first met for CR/PR (whichever is first recorded) until the first date that progressive disease is objectively documented (taking as reference for PD the smallest measurements recorded on study), subsequent anti-cancer therapy (including tumor-directed radiotherapy and tumor-directed surgery) is initiated, or the participant dies, whichever occurs first.

The duration of overall complete response is measured from the time measurement criteria are first met for CR until the first date that progressive disease is objectively documented, subsequent anti-cancer therapy (including tumor-directed radiotherapy and tumor-directed surgery) is initiated, or the participant dies, whichever occurs first.

4.3.2 Duration of Stable Disease

If SD is the best overall response, the duration of SD is measured from the date of randomization until the criteria for progression are met, taking as reference the smallest sum on study (if the baseline sum is the smallest, this is the reference for calculation of PD), subsequent anti-cancer therapy (including tumor-directed radiotherapy and tumor-directed surgery) is initiated, or the participant dies, whichever occurs first..

Reference: Eisenhauer EA, Therasse P, Bogaerts J, et al. New Response Evaluation Criteria in Solid Tumours: Revised RECIST Guideline (Version 1.1). Eur J Cancer 2009;45:228-247.

APPENDIX 9 INDUCERS AND INHIBITORS OF CYP3A4

Strong Inducers and Inhibitors of CYP3A4
CYP3A4 Inducers
Phenytoin
Carbamazepine
Rifampin
Rifabutin
Rifapentine
Phenobarbital
Dexamethasone
CYP3A4 Inhibitors
Ketoconazole
Itraconazole
Voriconazole
Clarithromycin
Erythromycin
Telithromycin
Nefazodone
Saquinavir
Ritonavir
Atazanavir
Indinavir
Nelfinavir

Notes: The above list is not exhaustive. Please refer to the drug interaction tables at the link below for lists of substrates, inducers, and inhibitors of selected CYP450 isozyme pathways. (Http://medicine.iupui.edu/clinpharm/ddis/table.aspx).

Grapefruit, grapefruit juice and other foods that are known to inhibit CYP3A4 activity should be avoided during treatment.

St. John's Wort (*Hypericum perforatum*) is known to be an inducer of CYP3A4 and should be avoided during treatment.

APPENDIX 10 SUPPORTIVE CARE GUIDELINES FOR THE MANAGEMENT OF HAND FOOT SYNDROME (HFS)

Supportive Care Guidelines for the Management of Hand Foot Syndrome (HFS)

Management of HFS can begin before any symptoms occur. Several prophylactic measures may be taken to prevent or reduce the severity of HFS. Before therapy with begins, a full-body skin exam should be performed, with a special emphasis on hyperkeratotic areas on palms and soles and any deformities. Patients can receive a pedicure, using properly sterilized utensils, to remove any preexisting hyperkeratotic areas or calluses that may predispose them to developing HFSR. Patients should be advised to reduce the exposure of their hands and feet to hot water, either through dishwashing or hot baths and showers, because this is believed to exacerbate symptoms, and patients frequently report symptomatic relief with cold water.

Before initiating treatment:

- Check condition of hands and feet
- Suggest a manicure/pedicure, when indicated
- Recommend pumice stone use for callus or 'rough spot' removal

During treatment:

- Avoid pressure points
- Avoid items that rub, pinch, or create friction
- Apply non-urea based skin-hydrating creams liberally
- Keratolytic creams: Use sparingly and only to affected (hyperkeratotic) areas.

Urea-based creams:

- Salicylic acid 6%
- Alpha hydroxy acid (AHA) based creams
- Concentrations of approximately 5-8% provide gentle chemical exfoliation
- Apply liberally two times each day

Topical analgesics like lidocaine 2% should be considered for pain control

Topical corticosteroids like clobetasol 0.05% should be considered for patients with grade 2 or 3 hand-foot skin reaction. Avoid systemic steroids.

Cushions:

- Protect tender areas
- Use socks/gloves to cover moisturizing creams
- Wear well-padded footwear
- Use insole cushions or inserts (e.g., silicon, gel)
- Foot soaks with tepid water and Epsom salts

APPENDIX 11 SUPPORTIVE CARE GUIDELINES FOR THE MANAGEMENT OF DIARRHEA FOR CABOZANTINIB

Participants should be instructed to notify their physician immediately at the first signs of poorly formed or loose stool or an increased frequency of bowel movements. Guidelines for the evaluation and management of diarrhea are shown in Table 1.

Administration of antidiarrheal/antimotility agents is recommended at the first sign of diarrhea as initial management. Some participants may require concomitant treatment with more than 1 antidiarrheal agent. When therapy with antidiarrheal agent does not control diarrhea to tolerable levels, cabozantinib should be temporarily interrupted or dose reduced. When the diarrhea is controlled, retreatment with cabozantinib may be acceptable per investigator decision.

In addition, general supportive measures should be implemented such as continuous oral isotonic hydration, correction of fluid and electrolyte abnormalities, small frequent meals, and stopping lactose-containing products, high-fat meals, and alcohol.

Recurrent or prolonged diarrhea can be associated with anal or perianal skin erosions which increase the risk for anal abscesses, fistulas, or proctitis. Good hygiene should be emphasized. Regular examinations of the perianal region should be performed wherever diarrhea has occurred during treatment with cabozantinib. Infections of the perianal region should be treated per local guidelines.

Table 1: Guidelines for Management of Treatment-Emergent Diarrhea for Cabozantinib

Status	Management
Tolerable Grade 1-2 (duration < 48 h)	• Continue with study treatment and consider dose reduction • Initiate treatment with an antidiarrheal agent (e.g., loperamide 4 mg followed by 2 mg after each episode of diarrhea [maximum: 16 mg loperamide daily]) • Dietary modifications (e.g., small lactose-free meals, bananas and rice) • Intake of isotonic fluids 1-1.5 L/day) • Re-assess after 24 hours: - Diarrhea resolving to baseline bowel habits: gradually add solid foods and discontinue or decrease antidiarrheal treatment after 12 h diarrhea-free interval - Diarrhea not resolving: Continue/resume antidiarrheal treatment
Intolerable Grade 2, Grade 2 > 48h, or ≥ Grade 3	• Interrupt study treatment • Ask participant to attend clinic • Rule out infection (e.g., stool sample for culture) - Administer antibiotics as needed, (e.g., if fever or Grade 3-4 neutropenia persists > 24 h) • Administer fluids 1-1.5L/day orally or IV, as appropriate for hydration or to correct electrolyte abnormalities • For Grade 3-4 or complicated lower grade diarrhea consider hospitalization and IV hydration • Re-assess after 24 hour • Diarrhea resolving to baseline bowel habits or Grade ≤ 1: - Consider restarting study treatment at reduced dose • Diarrhea not resolving: - Start and or continue antidiarrheal treatment (e.g., loperamide 4 mg followed by 2 mg after each episode of diarrhea [maximum 16 mg loperamide per day - Consider starting second line antidiarrheal or referral to gastroenterologist

APPENDIX 12 COUNTRY SPECIFIC REQUIREMENTS

1 COUNTRY SPECIFIC REQUIREMENTS

1.1 Germany

Original language	Country-specific language
Section 2 Schedule of Activities, Table 2-1, Screening Procedural Outline, Laboratory Tests	Add "HIV" to the list of laboratory tests
Section 6.2 Exclusion Criteria, Exclusion criterion 1 g)	"Known history of testing positive for human immunodeficiency virus (HIV) or known acquired immunodeficiency syndrome (AIDS)"to be replaced with "Positive test for HIV".

Protocol Number: CA2099ER
IND Number: 134,374
EUDRACT Number 2017-000759-20
Date 08-Mar-2017
Revised Date: 03-May-2019

Clinical Protocol CA2099ER

A Phase 3, Randomized, Open-Label Study of Nivolumab Combined with Cabozantinib versus Sunitinib in Participants with Previously Untreated Advanced or Metastatic Renal Cell Carcinoma

(CheckMate 9ER: CHECKpoint pathway and nivoluMAb clinical Trial Evaluation 9ER)

Revised Protocol Number: 02
Incorporates Administrative Letter(s): 02

Study Director/ Medical Monitor

24-hr Emergency Telephone Number
USA: 1-866-470-2267
International: +1-248-844-7390

Bristol-Myers Squibb Research and Development
3401 Princeton Pike
Lawrence Township, NJ 08648

Replace all previous version(s) of the protocol with this revised protocol and please provide a copy of this revised protocol to all study personnel under your supervision, and archive the previous versions.

DOCUMENT HISTORY

Document	Date of Issue	Summary of Change
Revised Protocol 02	03-May-2019	Major Changes: • Revised protocol 02 adjusts the timing of the PFS and OS interim analyses with modified hypothesized OS hazard ratio (HR). The number of randomized participants is increased. • The interim analysis for ORR is removed, resulting in revised overall alpha for PFS and OS endpoints. • No change in eligibility or study procedure. • Clinical data for nivolumuab + ipilimumab in renal cell carcinoma (RCC) has been updated. Other changes include more detail on PRO measures and updates to align with BMS standards for the nivolumab program.
Administrative Letter 02	08-Feb-2018	Clarification to the sites on the implementation of CA2099ER Global Revised Protocol 01 which stopped further randomization into Arm B, and the timing of IRB approval and impact on randomization. Protocol text was not changed.
Revised Protocol 01	18-Dec-2017	Primary revisions: (i) To stop enrollment into Arm B (nivolumab, ipilimumab and cabozantinib triplet) and (ii) to include favorable risk participants (capped at 25%) in the primary data analysis. Secondary items include: (i) to add a Data Monitoring Committee review after 30 participants are treated for 6 weeks, (ii) to adjust, clarify and add exclusion criteria, (iii) to add treatment restrictions, (iv) to clarify criteria associated with hemorrhage with regard to resuming treatment, (v) to specify an additional precaution when sunitinib dosing is resumed, and (vi) to apply newly updated Sponsor standards for nivolumab clinical protocols. Tertiary items include (i) incorporation of Administrative Letter 01 and (ii) correction of typographical and grammatical errors.
Administrative Letter 01	20-Jul-2017	To notify of a change of Medical Monitor and Study Director. To correct typographical errors and resolve inconsistencies found in the original protocol.
Original Protocol	08-Mar-2017	Not applicable

OVERALL RATIONALE FOR REVISED PROTOCOL 02:

Given that both the CheckMate-214[1] and KeyNote-426[2] studies have shown clinically and statistically significant improvements in overall survival (OS) of immunotherapy-containing combinations compared with single agent sunitinib in first-line clear cell metastatic RCC with OS hazard ratio (HR)=0.63 and 0.53, respectively, the CA209-9ER statistical analysis plan is being updated to improve OS assessment. To be more consistent with emerging data, the hypothesized OS HR is being updated to a more clinically meaningful 0.70 and the power to detect a difference in OS is being increased, with less required overall clinical follow-up time. The statistical plan is updated to reflect increased overall enrollment and accrual period. There has been no change to the study endpoints although the interim analysis for ORR is being removed with the corresponding 0.001 alpha reallocated to the primary PFS analysis. Patient-reported outcome measures and analyses are updated for clarity. Additional changes reflect alignment of the protocol with BMS standards for the nivolumab program.

SUMMARY OF KEY CHANGES FOR REVISED PROTOCOL 02

Section Number & Title	Description of Change	Brief Rationale
Synopsis, Exclusion Criteria 6.2 Exclusion Criteria (2h)	Added exclusion for live/attenuated vaccine	Reflects the new required standard language for BMS nivolumab studies
Synopsis, Objectives and Endpoints Table 4.1 Objectives and Endpoints Table 10.3.1-1 Efficacy Analyses	Added text to statistical analysis description of PFS	Clarified that those participants who had subsequent therapy before death will be censored
Synopsis, Overall Design 5.1 Overall Design	The total number of randomized participants is changed from approximately 580 to 638 participants (from 290 to 319 per arm for Arm A and C) and from 434 to 478 for intermediate/poor risk participants	The enrollment target is updated for further improvement of OS power
Synopsis, Number of Participants 5.2 Number of Participants	The total number of participants enrolled is changed from 774 to 850 as well as the number of randomized participants per arm A and C	See Overall Rationale above
Table 4-1 Objectives and Endpoints	Added row in exploratory analyses for assessment of PFS-2	To assess the effect of the study treatment beyond first progression
5.1.1 Data Monitoring Committee and Other External Committees 10.1 Sample Size Determination 10.3.4. Interim Analyses	Deleted ORR endpoint Deleted text and table on ORR endpoint Deleted first paragraph in 10.3.4	The interim analysis for ORR is being removed

SUMMARY OF KEY CHANGES FOR REVISED PROTOCOL 02		
Section Number & Title	**Description of Change**	**Brief Rationale**
6.2 Exclusion Criteria (5a) 8.1 Discontinuation from Study Treatment	Added "only in countries where local regulations permit"	Provides clarification
7.7.1 Prohibited and/or Restricted Treatments	Added prohibition for live/attenuated vaccine	Reflects new requirement for BMS nivolumab studies
9.1.4.1 FKSI-19 9.1.4.2 EQ-5D-3L	Created new subsections and revised and added text for FKSI-19 and EQ-5D-3L	To provide more detail on the properties of the PRO measures
10.1 Sample Size Determination Table 10.1-1 Summary of Sample Size Parameters and Schedule of Analyses Table 10.3.1-1 Efficacy Analyses 10.3.4. Interim Analyses	Removed early assessment of ORR, so overall alpha for PFS and OS endpoints is revised from 0.049 to 0.05 Accrual duration (and accruals per months) are updated from 15 months to 19 months Added requirements for minimal follow-up time of all randomized participant and minimal numbers of events required for final PFS and the 2 interim OS analyses Revised OS hypothesized HR assumption from 0.76 to 0.70. The overall power for OS is therefore increased to 80% Accordingly, the accruals, the number of events, HRs, the timing of analyses, the power for PFS and OS are updated in the correspondent sections and table 10.3.1-1 Deleted first paragraph of 10.3.4	The hypothesized HR for OS is updated based on emerging evidence of OS benefits observed in immunotherapy containing combinations. The power for OS as the key secondary endpoint is increased in line with the overall rationale
10.3.3.1 Outcomes Research Analyses Table 10.3.3.1-1 Thresholds Values for Change Scores Judged to be Important to Patients	Added new subsection and table	To provide more detail on analyses to be performed and the clinically meaningful scores to be used
Appendix 2 Study Governance Consideration	Slightly modified definition of "serious breach"; provided additional criteria for the CSR Signatory Investigator; added publication policy	Reflects the new required standard language for BMS nivolumab studies
Appendix 4 Women of Childbearing Potential Definitions and Methods of Contraception	Revised section on Highly Effective Methods That Are User Independent	Reflects the new required standard language for BMS nivolumab studies

[1]Motzer RJ, Tannir NM, McDermott DF, et al. Nivolumab plus Ipilimumab versus Sunitinib in Advanced Renal-Cell Carcinoma. N Engl J Med 2018;378(14):1277-1290.

[2] Rini BI, Plimack ER, Stus V, et al. Pembrolizumab plus Axitinib versus Sunitinib for Advanced Renal-Cell Carcinoma. N Engl J Med 2019;380:1116-1127.

TABLE OF CONTENTS

1 SYNOPSIS

Protocol Title: A Phase 3, Randomized, Open-Label Study of Nivolumab Combined with Cabozantinib versus Sunitinib in Participants with Previously Untreated Advanced or Metastatic Renal Cell Carcinoma

Study phase: 3

Rationale: Although multiple agents are approved as monotherapies for the treatment of patients with metastatic renal cell carcinoma (mRCC), the testing of combination therapies, in particular, treatment with immune-checkpoint inhibitors in combination with tyrosine kinase inhibitors (TKIs) has not been fully explored. While single agent therapies have improved outcomes, ongoing drug resistance and disease progression demonstrate an urgent need to find more effective therapies for mRCC patients. The top priorities in treating RCC continue to be improving progression free survival (PFS) and overall survival (OS), management of toxicities, and a better understanding of biomarkers. The CA2099ER trial will include previously untreated participants with mRCC and uses the well-characterized immune checkpoint inhibitor nivolumab in combination with cabozantinib, a known standard-of-care in previously treated mRCC participants. Nivolumab combined with cabozantinib may be an important step forward in evaluating combination regimens which could potentially optimize the management of previously untreated participants with mRCC.

Study Population: Male and female participants ≥ 18 years or the age of majority with previously untreated, advanced or metastatic renal cell carcinoma (RCC).

The following list contains key eligibility criteria only. For full list of eligibility criteria please see Section 6.

Key Inclusion Criteria

- Histological confirmation of RCC with a clear-cell component, including participants who may also have sarcomatoid features
- Advanced (not amenable to curative surgery or radiation therapy) or metastatic (American Joint Committee on Cancer [AJCC] Stage IV) RCC
- No prior systemic therapy for RCC with the following exception:
    - One prior adjuvant or neoadjuvant therapy for completely resectable RCC if such therapy did not include an agent that targets vascular endothelial growth factor (VEGF) or VEGF receptors and if recurrence occurred at least 6 months after the last dose of adjuvant or neoadjuvant therapy
- Karnofsky Performance Status (KPS) grade ≥ 70%
- Measurable disease as per RECIST v1.1 per investigator
- Tumor tissue, preferably obtained within 3 months but no more than 12 months prior to enrollment, with an associated pathology report, must be received by the central laboratory during screening for determination of PD-L1 expression. In order to be randomized, a participant must be classified as PD-L1 expression ≥ 1%, PD-L1 expression < 1%, or PD-L1 expression indeterminate.

- Participants with favorable, intermediate, and poor risk categories will be eligible for the study. Participants must be categorized according to favorable versus intermediate versus poor risk status at registration as per International Metastatic RCC Database Consortium (IMDC) criteria.
- Negative pregnancy test and able to meet protocol-specified reproductive requirements

Key Exclusion Criteria

- Any active central nervous system (CNS) metastases. Participants with treated, stable CNS metastases for at least 1 month are eligible
- Any tumor invading the superior vena cava (SVC), other major blood vessels, or GI tract; any evidence of endotracheal or endobronchial tumor
- Prior systemic treatment with VEGF, MET, AXL, KIT or RET targeted therapy (including, but not limited to, sunitinib, pazopanib, axitinib, tivozanib, sorafenib, lenvatinib, bevacizumab, and cabozantinib)
- Prior treatment with an anti-PD-1, anti-PD-L1, anti-PD-L2, anti-CD137, or anti-CTLA-4 antibody, or any other antibody or drug specifically targeting T-cell co-stimulation or checkpoint pathways
- Any active, known or suspected autoimmune disease or any condition requiring systemic treatment with corticosteroids (> 10 mg daily prednisone equivalent) or other immunosuppressive medications within 14 days of randomization
- Uncontrolled adrenal insufficiency
- Poorly controlled hypertension despite antihypertensive therapy
- History of unstable angina, myocardial infarction, symptomatic peripheral vascular disease, congestive heart failure (CHF, Class III or IV as defined by the New York Heart Association [NYHA]), or cerebrovascular accident (CVA)
- Deep vein thrombosis (DVT) or pulmonary embolism (PE) unless adequately treated with low molecular weight heparin (LMWH)
- Any unstable cardiac arrhythmia; prolonged QTcF > 450 msec for males and > 470 msec for females
- Serious, non-healing wound or ulcer; evidence of active bleeding or bleeding susceptibility; history of abdominal fistula, gastrointestinal perforation, intra-abdominal abscess, bowel or gastric outlet obstruction
- Concomitant strong CYP3A4 inducers or inhibitors within 14 days prior to randomization
- Ejection fraction ≤ 50% on screening echocardiogram or multigated acquisition scan (MUGA)
- Major surgery less than 6 weeks, nephrectomy less than 4 weeks, prior to randomization, with complete wound healing and no ongoing post-operative complications.
- Participants who have received a live/attenuated vaccine within 30 days of first treatment.

Objectives and Endpoints:

Objective	Endpoint
Primary	
To compare progression-free survival (PFS) per blinded independent central review (BICR) of nivolumab plus cabozantinib (Arm A, doublet) with sunitinib (Arm C) in all randomized participants.	The primary endpoint of this study is to compare the PFS per BICR of Arm A versus Arm C in all randomized participants. PFS is defined as the time between the date of randomization and the first date of the documented progression, or death due to any cause, whichever occurs first. Participants who die without a reported progression (and die without start of subsequent anti-cancer therapy) will be considered to have progressed on the date of their death. Participants who did not progress or die will be censored on the date of their last evaluable tumor assessment on or prior to initiation of subsequent anti-cancer therapy. Participants who did not have any on study tumor assessments and did not die will be censored on their date of randomization. Participants who started anti-cancer therapy without a prior reported progression will be censored on the date of their last evaluable tumor assessment on or prior to the initiation of first subsequent anti-cancer therapy.
Secondary	
To compare overall survival (OS) of Arm A with Arm C in all randomized participants.	The first secondary endpoint is to compare the OS of Arm A versus Arm C in all randomized participants. OS is defined as the time between the date of randomization and the date of death due to any cause. A participant who has not died will be censored at the last known alive date.
To evaluate the objective response rate (ORR) per BICR in-all randomized participants.	The second secondary endpoint is to describe ORR per BICR in all randomized participants. ORR is defined as the proportion of randomized participants who achieve a best response of complete response (CR) or partial response (PR) using the RECIST 1.1 criteria. Best overall response (BOR) is defined as the best response designation recorded between the date of randomization and the date of objectively documented progression per RECIST 1.1 or the date of subsequent therapy (including tumor-directed radiotherapy and tumor-directed surgery), whichever occurs first. For participants without documented progression or subsequent therapy, all available response designations will contribute to the BOR assessment. Duration of response (DOR) is defined as the time between the date of first confirmed documented response (CR or PR) to the date of first documented tumor progression (per RECIST 1.1) or death due to any cause, whichever occurs first. Participants who neither progress nor die will be censored on the date of their last tumor assessment. Responders who started anti-cancer therapy without a prior reported progression will be censored on the date of their last evaluable tumor assessment prior to the initiation of first subsequent anti-cancer therapy. Time to response (TTR) is defined as the time from randomization to the date of the first confirmed documented response (CR or PR), as assessed by BICR. DOR and TTR will be evaluated for responders (CR or PR) only.
To assess overall safety and tolerability in all treated participants.	As measured by the incidence of adverse events (AEs), serious adverse events (SAEs), AEs leading to discontinuation, deaths, laboratory abnormalities and changes from baseline.

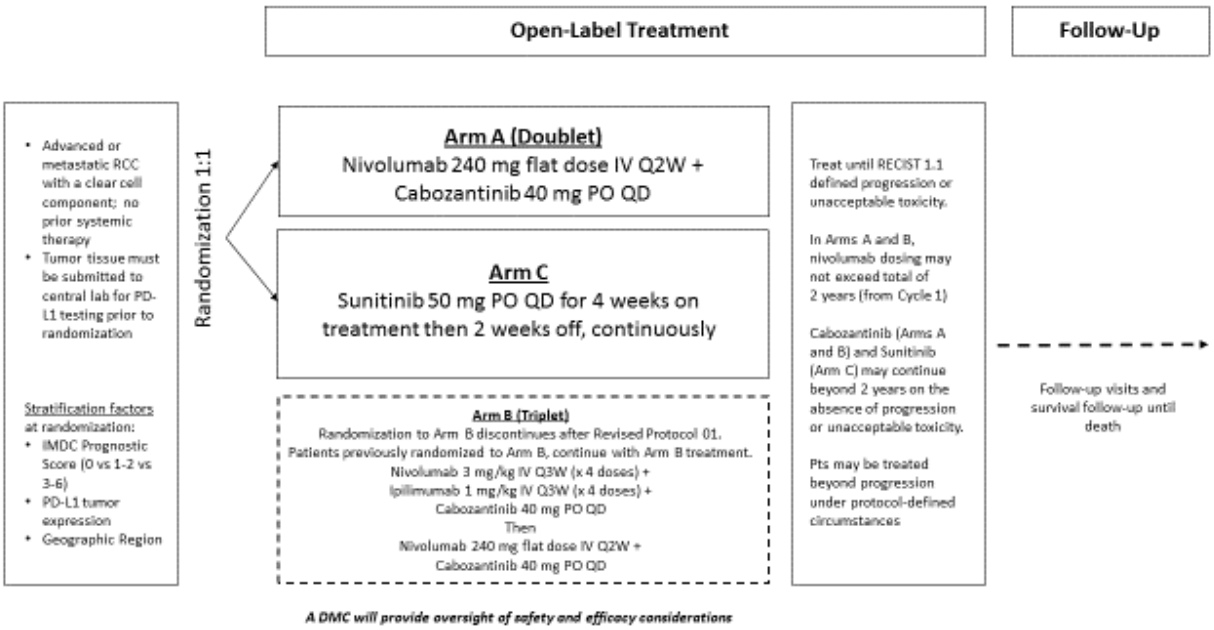
Overall Design:

CA2099ER Global Revised Protocol 01

Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B. Participants previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinical planned events, per protocol.

This is an open label, randomized trial of nivolumab combined with cabozantinib (doublet regimen) versus sunitinib in participants with previously untreated (first line) advanced or metastatic RCC. Participants will be randomized between Arm A and Arm C in a 1:1 ratio with approximately 638 participants (319 per Arms A and C, respectively), and capped at 25% favorable risk participants. Participants will be stratified for randomization by IMDC prognostic score (0 [favorable risk] versus 1-2 [intermediate risk] versus 3-6 [poor risk]), PD-L1 tumor expression (≥ 1% versus < 1% or indeterminate), and region (US/Canada/Western Europe/Northern Europe versus rest of the world [ROW]).

The study design schematic is presented below.



Abbreviations: DMC= data monitoring committee; IMDC= International Metastatic Renal Cell Carcinoma Database Consortium; IV=intravenous;; PD-L1= programmed death-ligand 1; PO= orally by mouth; Pts=participants; Q2W=every 2 weeks; Q3W=every 3 weeks; QD= once daily; RCC=renal cell carcinoma.

Arm B is shown in the design schematic because after implementation of Revised Protocol 01, there will be patients on Arm B treatment from the original protocol version 08 Mar 2017.

Number of Participants: Approximately 850 participants will enroll in order to randomize 638 participants into Arms A and C (319 per Arm A and per Arm C) capped at 25% favorable risk participants.

Treatment Arms and Duration:

CA2099ER Global Revised Protocol 01

Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B. Participants previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinical planned events, per protocol.

- Arm A (Doublet): Nivolumab 240 mg flat dose intravenously (IV) every 2 weeks (Q2W) + Cabozantinib 40 mg orally by mouth (PO) once daily (QD)
    - Nivolumab treatment until disease progression or unacceptable toxicity with maximum treatment of 2 years
    - Cabozantinib treatment until disease progression or unacceptable toxicity
- Arm C: Sunitinib 50 mg PO QD for 4 weeks, followed by 2 weeks off-treatment, per cycle. Cycles to be continued until progression or unacceptable toxicity

Note - Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B. Participants previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinical planned events, per protocol.

- Arm B (Triplet): Nivolumab 3mg/kg IV Q3W x 4 doses + Ipilimumab 1 mg/kg IV Q3W x 4 doses + Cabozantinib 40 mg PO QD
    - Then, Nivolumab 240 mg flat dose IV Q2W + Cabozantinib 40 mg PO QD
    - Nivolumab to be continued until disease progression or unacceptable toxicity with maximum treatment of 2 years from the start of first dose in Cycle 1
    - Cabozantinib until disease progression or unacceptable toxicity

Refer to Section 7.1 Treatments Administered for additional details.

Study treatment:

Study Drugs for CA2099ER

Medication	Potency	IP/ Non-IP
BMS-936558-01 (Nivolumab) Solution for Injection	100 mg (10 mg/mL)	IP
Ipilimumab Solution for Injection*	200 mg (5 mg/mL)	IP
Cabozantinib Tablet	20 mg	IP
Sunitinib Malate Capsule	12.5 mg	IP

Abbreviation: IP=investigational product

*Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B. Participants previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinical planned events, per protocol.

2 SCHEDULE OF ACTIVITIES

Table 2-1: Screening Procedural Outline (CA2099ER)

Procedure	Screening Visit[a]	Notes
Eligibility Assessments		
Informed Consent	X	Register in Interactive Response Technology (IRT) system to obtain participant number. The participant should sign the Informed consent prior to any study related assessment is performed.
Inclusion/Exclusion Criteria	X	Must be confirmed prior to randomization.
Medical History	X	
International Metastatic RCC Database Consortium (IMDC) Prognostic Score	X	See Appendix 6.
Tumor tissue sample (for stratification by PD-L1 tumor expression)	X	Tumor tissue (preferably obtained within 3 months but no more than 12 months prior to enrollment, with an associated pathology report) will be collected. Formalin-fixed paraffin-embedded (FFPE) block or 20 unstained slides: a minimum of 10 slides will be acceptable if tumor tissue is limited. See Section 9.8.2. Central lab will determine PD-L1 tumor expression. Participants must have an evaluable PD-L1 result from the central lab in order to be randomized.
Safety Assessments		
Full Physical Examination, Measurements, Vital Signs, and Performance Status	X	Height, weight, Karnofsky Performance Status (KPS) (Appendix 7), BP, HR, RR, and temperature within 14 days prior to randomization.
Assessment of Signs and Symptoms	X	Within 14 days prior to randomization.
Review of Concomitant Medications	X	Within 14 days prior to randomization.
Serious Adverse Events Assessment	X	Serious Adverse Events from time of consent. See Section 9.2
Electrocardiogram (ECG)	X	Within 28 days prior to randomization. Fridericia corrected QT (QTcF) required. If any time there is an increase in QTcF interval to an absolute value > 500 msec, 2 additional ECGs must be performed each with intervals approximately 3 minutes apart.

Table 2-1: Screening Procedural Outline (CA2099ER)

Procedure	Screening Visit[a]	Notes
Cardiac Ejection Fraction (via Echocardiogram or MUGA)	X	Within 28 days prior to randomization.
Laboratory Tests (includes blood and urine samples)	X	See Section 9.4.1 for additional details on tests required. To be completed locally at each site. Must be performed within 14 days prior to randomization. • CBC w/differential • PT/INR, PTT • Chemistry panel (includes AST, ALT, total bilirubin, ALP, LDH, creatinine, BUN, glucose, albumin, Na, K, Cl, Ca (also corrected), P, Mg, amylase, lipase) • Thyroid panel (includes TSH with free T3 and free T4) • Hepatitis B/C (HBVsAg, HCV antibody or HCV RNA) • HIV if mandated locally (sites in Germany, see Appendix 12) • Urine protein and urine creatinine (for urine protein/creatinine ratio [UPCR]). If UPCR ≥ 1.0, obtain 24 hour urine protein.
Pregnancy Test	X	WOCBP only. Serum or urine pregnancy test (minimum sensitivity 25 IU/L or equivalent units of HCG) to be done at screening visit and within 24 hours of first dose of study therapy.
Follicle Stimulating Hormone (FSH)	X	For women under the age of 55 years to confirm menopause as needed.
Efficacy Assessments		
Baseline Tumor Assessments	X	CT/MRI of the chest, abdomen, pelvis, brain, and all known sites of disease, performed within 28 days prior to randomization. All scans need to be submitted for blinded independent central review (BICR). See Section 9.1.

Abbreviations: For abbreviations on lab tests refer back to Section 9.4.1.

[a] Some of the assessments referred to in this section may not be captured as data in the eCRF. They are intended to be used as safety monitoring by the treating physician. Additional testing or assessments may be performed as clinically necessary or where required by institutional or local regulations.

Table 2-2: On Treatment Procedural Outline for Arm A Nivolumab Combined with Cabozantinib, Doublet (CA2099ER)

Procedure	Cycle 1, Day 1 Cycle=2 wks	C2 and subsequent visits[a] Each Cycle=2 wks	Notes
Safety Assessments			
Targeted Physical Examination, Vital Signs, Performance Status	X	X	Weight, BP, HR, RR, temperature, and Karnofsky Performance Status (KPS) (Appendix 7). Performance Status to be performed within 72 hours prior to dosing.
Assessment of Signs and Symptoms	X	X	
Adverse Events and Serious Adverse Events Assessment	X	X	Record at each visit. Participants will be followed for drug-related toxicities until these toxicities resolve, return to baseline, or are deemed irreversible. All adverse events will be documented for a minimum of 100 days after last dose (Appendix 3).
Concomitant Medication Review	X	X	Record at each visit.
Electrocardiogram (ECG)	X	X (See notes)	Fridericia corrected QT (OTcF) required. Only for Cycles 1, 4, 7, then every 6 cycles (ie, Cycles 13, 19, 25, etc). If any time there is an increase in QTcF interval to an absolute value > 500 msec, 2 additional ECGs must be performed each with intervals approximately 3 minutes apart.
Laboratory Tests (includes blood and urine samples)	X	X (See notes)	See 9.4.1 for additional details on tests required. Laboratory tests do not need to be repeated at C1D1 if performed within 14 days prior to first dose. After C1D1, within 72 hours prior to dosing: • CBC w/differential at every cycle • Chemistry panel at every cycle (includes AST, ALT, total bilirubin, ALP, LDH, creatinine, BUN, glucose, albumin, Na, K, Cl, Ca, P, Mg) • Amylase and lipase to be done for Cycles, 1, 2, 4, 5, 7, and then every 3 cycles (ie, Cycles 10, 13, 16, etc) • Thyroid panel (includes TSH with reflexive free T3 and free T4) for Cycles 1, 2, 4, 5, 7, and then every 3 cycles (ie, Cycles 10, 13, 16, etc) • Urine protein and urine creatinine (for UPCR, preferred) or urine dipstick for protein every 3 cycles (ie, Cycles 1, 4, 7, 10, etc).

Table 2-2: On Treatment Procedural Outline for Arm A Nivolumab Combined with Cabozantinib, Doublet (CA2099ER)

Procedure	Cycle 1, Day 1 Cycle=2 wks	C2 and subsequent visits[a] Each Cycle=2 wks	Notes
Pregnancy Test	X	X	Within 24 hours prior to the initial administration of study drug, then every 4 weeks ± 7 days. Serum or Urine. WOCBP only.
Efficacy Assessments			
Tumor Assessments	First tumor assessment post-baseline should be performed Week 12 (± 7 days) following randomization. Use same imaging method as was used at baseline. Subsequent tumor assessments should occur at every 6 weeks (± 7 days) until Week 60, then every 12 weeks (± 14 days) until radiographic progression, assessed by the investigator and confirmed by the BICR.		CT/MRI of the chest, abdomen, pelvis, and all known sites of disease. Tumor assessments should be performed at the specified time points regardless of dosing delays. See Section 9.1 for additional details. Treatment Beyond Progression A radiographic assessment/ scan should be performed within 6 weeks of initial investigator-assessed progression. Refer to Section 8.1.4 for tumor assessment associated with Treatment beyond progression.
Pharmacokinetic (PK)/ Immunogenicity Assessments			
PK blood samples	X	X (See notes)	For details on sampling time points see Table 9.5-1.
Immunogenicity blood samples	X	X (See notes)	For details on sampling time points see Table 9.5-1.
Exploratory Biomarker Assessments			
Whole Blood (DNA) for Genotyping	X		Only prior to dose at Cycle 1. See Sections 9.8.1 and 9.8.6.
Serum Biomarker	X	X (See notes)	Prior to dosing. At Cycles 1, 2, and 4. See Sections 9.8.1 and 9.8.3.

Table 2-2: On Treatment Procedural Outline for Arm A Nivolumab Combined with Cabozantinib, Doublet (CA2099ER)

Procedure	Cycle 1, Day 1 Cycle=2 wks	C2 and subsequent visits[a] Each Cycle=2 wks	Notes
Plasma Biomarkers	X	X (See notes)	Prior to dosing. At Cycles 1, 2, and 4. See Sections 9.8.1 and 9.8.3.
Myeloid Derived Suppressor Cells	X		Prior to dosing. At Cycle 1 only. See Section 9.8.4.
Peripheral blood mononuclear cells (PBMCs)	X	X (See notes)	Prior to dosing: At Cycles 1 and 4. See Sections 9.8.1 and 9.8.5
Tumor Tissue Sample	Every effort should be made to collect fresh tumor tissue sample if available upon disease progression.		If a biopsy has been performed and tissue is available, participants are requested to submit fresh tumor tissue upon disease progression for biomarker research. Tissue submission is optional and biopsy is not required by protocol. See Section 9.8.2. Treatment Beyond Progression A radiographic assessment/ scan should be performed within 6 weeks of initial investigator-assessed progression. Refer to Section 8.1.4 for tumor assessment associated with Treatment beyond progression.
Participant-Reported Outcomes			
Functional Assessment of Cancer Therapy - Kidney Symptom Index (FKSI-19) Questionnaire	X	X	Completed on Day 1 of each treatment cycle prior to any study-related procedures. See Section 9.1.4.
EuroQoL group's EQ-5D-3L Questionnaire	X	X	Completed on Day 1 of each treatment cycle prior to any study-related procedures. See Section 9.1.4
Health Care Resource Utilization	X	X	See Section 9.9.
Study Treatment			

Table 2-2: On Treatment Procedural Outline for Arm A Nivolumab Combined with Cabozantinib, Doublet (CA2099ER)

Procedure	Cycle 1, Day 1 Cycle=2 wks	C2 and subsequent visits[a] Each Cycle=2 wks	Notes
Randomize	X		Begins with call to Interactive Response Technology (IRT). Participants must have an evaluable PD-L1 result from the central lab in order to be randomized.
Administer Nivolumab and Cabozantinib	X	X	See Section 7. Dispense study treatment as appropriate. Day 1 Treatment must begin within 3 days (72 hours) of Randomization)

Abbreviations: C=cycle; D=day; wks= weeks. For abbreviations on lab tests refer back to Section 9.4.1.

Notes: Some of the assessments referred to in this section may not be captured as data in the eCRF. They are intended to be used as safety monitoring by the treating physician. Additional testing or assessments may be performed as clinically necessary or where required by institutional or local regulations.

[a] If a dose is delayed, the procedures scheduled for that same timepoint should also be delayed to coincide with when that timepoint's dosing actually occurs.

Table 2-3: On Treatment Procedural Outline for Arm B Nivolumab and Ipilimumab Combined with Cabozantinib, Triplet (CA2099ER)

CA2099ER Global Revised Protocol 01

Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B. Participants previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinical planned events, per protocol.

Procedure	Part 1		Part 2	Notes (Please note differences in cycle durations between Part 1 and Part 2)
	Cycle 1 (Each Cycle=3 wks)	C2-C4[a] (Each Cycle= 3 wks)	C5 and Subsequent visits [a] (Each Cycle=2 wks)	
Safety Assessments				
Targeted Physical Examination, Vital Signs, Performance Status	X	X	X	Weight, BP, HR, RR, temperature, and Karnofsky Performance Status (KPS) (Appendix 7). Performance Status to be performed within 72 hours prior to dosing.
Assessment of Signs and Symptoms	X	X	X	
Adverse Events and Serious Adverse Events Assessment	X	X	X	Record at each visit. Participants will be followed for drug-related toxicities until these toxicities resolve, return to baseline, or are deemed irreversible. All adverse events will be documented for a minimum of 100 days after last dose (Appendix 3).
Concomitant Medication Review	X	X	X	Record at each visit.
Electrocardiogram (ECG)	X	X (See notes)	X (See notes)	Fridericia corrected QT (QTcF) required. Only Cycles 1, 3, 5, then every 6 cycles (ie, Cycles 11, 17, 23, etc). If any time there is an increase in QTcF interval to an absolute value > 500 msec, 2 additional ECGs must be performed each with intervals approximately 3 minutes apart.
Laboratory Tests (includes blood and urine samples)	X	X	X (See notes)	See Section 9.4.1 for additional details on tests required. Laboratory tests do not need to be repeated at C1D1 if performed within 14 days prior to first dose. After C1D1, within 72 hours prior to dosing. • CBC w/differential at every cycle

Table 2-3: On Treatment Procedural Outline for Arm B Nivolumab and Ipilimumab Combined with Cabozantinib, Triplet (CA2099ER)

CA2099ER Global Revised Protocol 01

Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B. Participants previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinical planned events, per protocol.

Procedure	Part 1		Part 2	Notes (Please note differences in cycle durations between Part 1 and Part 2)
	Cycle 1 (Each Cycle=3 wks)	C2-C4[a] (Each Cycle= 3 wks)	C5 and Subsequent visits [a] (Each Cycle=2 wks)	
				• Chemistry panel at every cycle (includes AST, ALT, total bilirubin, ALP, LDH, creatinine, BUN, glucose, albumin, Na, K, Cl, Ca, P, Mg) • Amylase and lipase at every cycle until Cycle 5 then every 3 cycles (ie, Cycles 8, 11, 14, etc) • Thyroid panel (includes TSH with reflexive free T3 and free T4) at every cycle until Cycle 5, then every 3 cycles (ie, Cycles 8, 11, 14, etc) • Urine protein and urine creatinine (for UPCR, preferred) or urine dipstick for protein at Cycles 1, 3, and 5, then every 3 cycles (ie, Cycles 8, 11, 14, etc)
Pregnancy Test	X	X	X	Within 24 hours prior to the initial administration of study drug, then every 4 weeks ± 7 days. Serum or Urine. WOCBP only.
Efficacy Assessments				
Tumor Assessments	First tumor assessment post-baseline should be performed Week 12 (± 7 days) following randomization. Use same imaging method as was used at baseline. Subsequent tumor assessments should occur at every 6 weeks (± 7 days) until Week 60, then every 12 weeks (± 14 days) until radiographic progression, assessed by the investigator and confirmed by the BICR.			CT/MRI of the chest, abdomen, pelvis, and all known sites of disease. Tumor assessments should be performed at the specified time points regardless of dosing delays. See Section 9.1 for additional details. Treatment Beyond Progression A radiographic assessment/ scan should be performed within 6 weeks of initial investigator-assessed progression. Refer to Section 8.1.4 for tumor assessment associated with Treatment beyond progression.

Table 2-3: On Treatment Procedural Outline for Arm B Nivolumab and Ipilimumab Combined with Cabozantinib, Triplet (CA2099ER)

CA2099ER Global Revised Protocol 01

Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B. Participants previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinical planned events, per protocol.

Procedure	Part 1		Part 2	Notes (Please note differences in cycle durations between Part 1 and Part 2)
	Cycle 1 (Each Cycle=3 wks)	C2-C4[a] (Each Cycle= 3 wks)	C5 and Subsequent visits [a] (Each Cycle=2 wks)	
Pharmacokinetic (PK) /Immunogenicity Assessments				
PK blood samples	X	X (See notes)	X (See notes)	For details on sampling timepoints see Table 9.5-2.
Immunogenicity blood samples	X	X (See notes)	X (See notes)	For details on sampling timepoints see Table 9.5-2.
Exploratory Biomarker Assessments				
Whole Blood (DNA) for Genotyping	X			Only prior to dose at Cycle 1. See Sections 9.8.1 and 9.8.6.
Serum Biomarkers	X	X (See notes)		Prior to dosing. At Cycles 1, 2, and 3. See Sections 9.8.1 and 9.8.3.
Plasma Biomarkers	X	X (See notes)		Prior to dosing. At Cycles 1, 2, and 3. See Sections 9.8.1 and 9.8.3.
Myeloid Derived Suppressor Cells	X			Prior to dosing. At Cycle 1 only. See Section 9.8.4.
Peripheral blood mononuclear cells (PBMCs)	X	X (See notes)		Prior to dosing. At Cycles 1 and 3 only. See Sections 9.8.1 and 9.8.5.

Table 2-3: On Treatment Procedural Outline for Arm B Nivolumab and Ipilimumab Combined with Cabozantinib, Triplet (CA2099ER)

CA2099ER Global Revised Protocol 01

Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B. Participants previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinical planned events, per protocol.

Procedure	Part 1		Part 2	Notes (Please note differences in cycle durations between Part 1 and Part 2)
	Cycle 1 (Each Cycle=3 wks)	C2-C4[a] (Each Cycle= 3 wks)	C5 and Subsequent visits [a] (Each Cycle=2 wks)	
Tumor Tissue Sample	Every effort should be made to collect fresh tumor tissue sample if available upon disease progression.			If a biopsy has been performed and tissue is available, participants are requested to submit fresh tumor tissue upon disease progression for biomarker research. See Section 9.8.2. Tissue submission is optional and biopsy is not required by protocol. Treatment Beyond Progression A radiographic assessment/ scan should be performed within 6 weeks of initial investigator-assessed progression. Refer to Section 8.1.4 for tumor assessment associated with Treatment beyond progression.
Participant-Reported Outcomes				
Functional Assessment of Cancer Therapy - Kidney Symptom Index (FKSI-19) Questionnaire	X	X	X	Completed on Day 1 of each treatment cycle prior to any study-related procedures. See Section 9.1.4.
EuroQoL group's EQ-5D-3L Questionnaire	X	X	X	Completed on Day 1 of each treatment cycle prior to any study-related procedures. See Section 9.1.4.
Health Care Resource Utilization	X	X	X	See Section 9.9.
Study Treatment				

Table 2-3: On Treatment Procedural Outline for Arm B Nivolumab and Ipilimumab Combined with Cabozantinib, Triplet (CA2099ER)

CA2099ER Global Revised Protocol 01

Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B. Participants previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinical planned events, per protocol.

Procedure	Part 1		Part 2	Notes (Please note differences in cycle durations between Part 1 and Part 2)
	Cycle 1 (Each Cycle=3 wks)	C2-C4[a] (Each Cycle= 3 wks)	C5 and Subsequent visits [a] (Each Cycle=2 wks)	
Randomize	X			Begins with call to Interactive Response Technology (IRT). Participants must have an evaluable PD-L1 result from the central lab in order to be randomized. Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B. Participants previously randomized to Arm B (Triplet) continue with Arm B treatment and continue with Arm B clinically planned events, per protocol.
Administer Nivolumab, Ipilimumab, and Cabozantinib	X	X		Cycles 1 to 4 are 3 week cycles. See Section 7. Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B. Participants previously randomized to Arm B (Triplet) continue with Arm B treatment and continue with Arm B clinically planned events, per protocol.
Administer Nivolumab and Cabozantinib			X	Day 1 Treatment must begin within 3 days (72 hours) of Randomization) Cycle 5 and subsequent cycles are 2 week cycles. See Section 7.
Dispense Study Treatment	X	X	X	See Section 7.

Abbreviations: C=cycle; D=day; wks= weeks. For abbreviations on lab tests refer back to Section 9.4.1.

Notes: Some of the assessments referred to in this section may not be captured as data in the eCRF. They are intended to be used as safety monitoring by the treating physician. Additional testing or assessments may be performed as clinically necessary or where required by institutional or local regulations.

[a] If a dose is delayed, the procedures scheduled for that same timepoint should also be delayed to coincide with when that timepoint's dosing actually occurs.

Table 2-4: On Treatment Procedural Outline for Arm C Sunitinib (CA2099ER)

Procedure	Cycle 1 Each Cycle=6wks		Cycle 2 [a] Each Cycle=6 wks		C3 and Subsequent visits[a] Each Cycle=6 wks		Notes
	Day 1	Day 22 (+/- 3 days)	Day 1	Day 22 (+/- 3 days)	Day 1	Day 22 (+/- 3 days)	
Safety Assessments							
Targeted Physical Examination, Vital Signs, Performance Status	X		X		X		Weight, BP, HR, RR, temperature, and Karnofsky Performance Status (KPS) (Appendix 7). Performance Status to be performed within 72 hours prior to dosing.
Assessment of Signs and Symptoms	X	X	X	X	X	X	
Adverse Events and Serious Adverse Events Assessment	X	X	X	X	X	X	Record at each visit. Participants will be followed for drug-related toxicities until these toxicities resolve, return to baseline, or are deemed irreversible. All adverse events will be documented for a minimum of 100 days after last dose (Appendix 3).
Concomitant Medication Review	X	X	X	X	X	X	Record at each visit
Electrocardiogram (ECG)	X		X		X (See notes)		Fridericia corrected QT (QTcF) required. Only Cycles 1, 2, and 3, then every 2 cycles (ie, Cycles 5, 7, 9, etc). If any time there is an increase in QTcF interval to an absolute value > 500 msec, 2 additional ECGs must be performed each with intervals approximately 3 minutes apart.
Laboratory Tests (includes blood and urine samples)	X	X (See notes)	X	X (See notes)	X	X (See notes)	See Section 9.4.1 for additional details on tests required. Laboratory tests do not need to be repeated at C1D1 if performed within 14 days prior to first dose. After C1D1, within 72 hours prior to specified dosing day: • CBC w/differential on Day 1 and Day 22 (+/- 3 days) of each cycle

Table 2-4: On Treatment Procedural Outline for Arm C Sunitinib (CA2099ER)

Procedure	Cycle 1 Each Cycle=6wks		Cycle 2 [a] Each Cycle=6 wks		C3 and Subsequent visits[a] Each Cycle=6 wks		Notes
	Day 1	Day 22 (+/- 3 days)	Day 1	Day 22 (+/- 3 days)	Day 1	Day 22 (+/- 3 days)	
							• Chemistry panel on Day 1 and Day 22 (+/- 3 days) of every cycle (includes AST, ALT, total bilirubin, ALP, LDH, creatinine, BUN, glucose, albumin, Na, K, Cl, Ca, P, Mg) • Amylase and lipase on Day 1 and Day 22 (+/- 3 days) of Cycle 1 and 2, then on Day 1 of Cycle 3 and all subsequent cycles. • Thyroid panel (includes TSH with reflexive free T3 and free T4) on Day 1 and Day 22 (+/- 3 days) of Cycle 1 and 2, then on Day 1 of Cycle 3 and all subsequent cycles. • Urine protein and urine creatinine (for UPCR, preferred) or urine dipstick for protein on Day 1 of every cycle
Pregnancy Test	X	X	X	X	X	X	Within 24 hours prior to the initial administration of study drug, then every 4 weeks ± 7 days. Serum or Urine. WOCBP only.
Efficacy Assessments							
Tumor Assessments	First tumor assessment post-baseline should be performed Week 12 (± 7 days) following randomization. Use same imaging method as was used at baseline. Subsequent tumor assessments should occur at every 6 weeks (± 7 days) until Week 60, then every 12 weeks (± 14 days) until radiographic progression, assessed by the investigator and confirmed by the BICR.						CT/MRI of the chest, abdomen, pelvis, and all known sites of disease. Tumor assessments should be performed at the specified time points regardless of dosing delays. See Section 9.1 for additional details. Treatment Beyond Progression A radiographic assessment/ scan should be performed within 6 weeks of initial investigator-assessed

Table 2-4: On Treatment Procedural Outline for Arm C Sunitinib (CA2099ER)

Procedure	Cycle 1 Each Cycle=6wks		Cycle 2 [a] Each Cycle=6 wks		C3 and Subsequent visits[a] Each Cycle=6 wks		Notes
	Day 1	Day 22 (+/- 3 days)	Day 1	Day 22 (+/- 3 days)	Day 1	Day 22 (+/- 3 days)	
							progression. Refer to Section 8.1.4 for tumor assessment associated with Treatment beyond progression.
Exploratory Biomarker Assessments							
Whole Blood (DNA) for Genotyping	X						Only prior to dose at Cycle 1. See Sections 9.8.1 and 9.8.6
Serum Biomarkers	X		X				Prior to dosing. At Cycles 1 and 2. See Sections 9.8.1 and 9.8.3.
Plasma Biomarkers	X		X				Prior to dosing. At Cycles 1 and 2. See Sections 9.8.1 and 9.8.3.
Myeloid Derived Suppressor Cells	X						Prior to dosing at Cycle 1 only. See Section 9.8.4.
Peripheral blood mononuclear cells (PBMCs)	X		X				Prior to dosing. At Cycles 1 and 2. See Sections 9.8.1 and 9.8.5.
Tumor Tissue Sample	Every effort should be made to collect fresh tumor tissue sample if available upon disease progression.						If a biopsy has been performed and tissue is available, participants are requested to submit fresh tumor tissue upon disease progression for biomarker research. See Section 9.8.2. Tissue submission is optional and biopsy is not required by protocol. Treatment Beyond Progression A radiographic assessment/ scan should be performed within 6 weeks of initial investigator-assessed progression. Refer to Section 8.1.4 for tumor assessment associated with Treatment beyond progression.

Table 2-4: On Treatment Procedural Outline for Arm C Sunitinib (CA2099ER)

Procedure	Cycle 1 Each Cycle=6wks		Cycle 2 [a] Each Cycle=6 wks		C3 and Subsequent visits[a] Each Cycle=6 wks		Notes
	Day 1	Day 22 (+/- 3 days)	Day 1	Day 22 (+/- 3 days)	Day 1	Day 22 (+/- 3 days)	
Participant-Reported Outcomes							
Functional Assessment of Cancer Therapy - Kidney Symptom Index (FKSI-19) Questionnaire	X		X		X		Completed on Day 1 of each treatment cycle prior to any study-related procedures. See Section 9.1.4.
EuroQoL group's EQ-5D-3L Questionnaire	X		X		X		Completed on Day 1 of each treatment cycle prior to any study-related procedures. See Section 9.1.4.
Health Care Resource Utilization	X		X		X		See Section 9.9.
Study Treatment							
Randomize	X						Begins with call to Interactive Response Technology (IRT). Participants must have an evaluable PD-L1 result from the central lab in order to be randomized.
Administer Sunitinib	X	X	X	X	X	X	Day 1 Treatment must begin within 3 days (72 hours) of Randomization) Each cycle will be 6 weeks where sunitinib will be administered for 4 weeks, then participants will be off treatment for 2 weeks. See Section 7.
Dispense Study Treatment	X		X		X		See Section 7.

Abbreviations: C=cycle; D=day; wks= weeks. For abbreviations on lab tests refer back to Section 9.4.1.

Notes: Some of the assessments referred to in this section may not be captured as data in the eCRF. They are intended to be used as safety monitoring by the treating physician. Additional testing or assessments may be performed as clinically necessary or where required by institutional or local regulations.

[a] If a dose is delayed, the procedures scheduled for that same timepoint should also be delayed to coincide with when that timepoint's dosing actually occurs.

Table 2-5: Follow-Up Procedural Outline for All Arms (CA2099ER)

Procedure	Safety Follow-up (Follow up Visit 1 (FU1) and Visit 2 (FU2)[a]	Survival Follow-up[b]	Notes
Safety Assessments			
Targeted Physical Examination, Vital Signs, Performance Status	X		Weight, BP, HR, RR, temperature, and Karnofsky Performance Status (KPS) (Appendix 7).
Assessment of Signs and Symptoms	X		
Adverse Events and Serious Adverse Events Assessment	X		Record at each visit. Participants will be followed for drug-related toxicities until these toxicities resolve, return to baseline, or are deemed irreversible. All adverse events will be documented for a minimum of 100 days after last dose (Appendix 3).
Concomitant Medication Review	X		Record at each visit
Electrocardiogram (ECG)	X		Fridericia corrected QT (QTcF) required. If any time there is an increase in QTcF interval to an absolute value > 500 msec, 2 additional ECGs must be performed each with intervals approximately 3 minutes apart.
Laboratory Tests (includes blood and urine samples)	FU1 -yes FU2 - if toxicities are present		See Section 9.4.1 for additional details on tests required. • CBC w/differential, PT/INR, and PTT • Chemistry panel (includes AST, ALT, total bilirubin, ALP, LDH, creatinine, BUN, glucose, albumin, Na, K, Cl, Ca, P, Mg, amylase, lipase) • Thyroid panel (includes TSH with reflexive free T3 and free T4) • Urine protein and urine creatinine (for UPCR, preferred) or urine dipstick for protein
Pregnancy Test	X		Serum or Urine. WOCBP only.
Efficacy Assessments			
Tumor Assessments	First tumor assessment post-baseline should be performed Week 12 (± 7 days) following randomization. Use same imaging method as was used at baseline.		Participants who discontinue study treatment without radiographic progression, confirmed by BICR, will continue tumor assessments according to the protocol specified schedule, even if new anti-tumor therapy has been initiated in the Follow-Up phase, until radiographic

Table 2-5: Follow-Up Procedural Outline for All Arms (CA2099ER)

Procedure	Safety Follow-up (Follow up Visit 1 (FU1) and Visit 2 (FU2)[a]	Survival Follow-up[b]	Notes
	Subsequent tumor assessments should occur at every 6 weeks (± 7 days) until Week 60, then every 12 weeks (± 14 days) until radiographic progression, assessed by the investigator and confirmed by the BICR.		progression has been assessed by the investigator and confirmed by BICR. CT/MRI of the chest, abdomen, pelvis, and all known sites of disease. See Section 9.1.2 for additional details.
Survival Status	X	X	During safety follow up and every 3 months (clinic visit or by telephone) during survival phase. Include documentation of subsequent chemotherapy. See Section 8.1.5.
Pharmacokinetic (PK)/ Immunogenicity Assessments			
PK blood samples	X (See notes)		Only for participants who were in Arm A (Nivolumab combined with cabozantinib, Doublet) or Arm B (Nivolumab and ipilimumab combined with cabozantinib, Triplet). For details on sampling timepoints see Table 9.5-1
Immunogenicity blood samples	X (See notes)		Only for participants who were in Arm A (Nivolumab combined with cabozantinib, Doublet) or Arm B (Nivolumab and ipilimumab combined with cabozantinib, Triplet). For details on sampling timepoints see Table 9.5-1
Participant-Reported Outcomes			
Functional Assessment of Cancer Therapy - Kidney Symptom Index (FKSI-19) Questionnaire	X		See Section 9.1.4
EuroQoL group's EQ-5D-3L Questionnaire	X	X	See Section 9.1.4.
Health Care Resource Utilization	X		See Section 9.9.

Abbreviations: C=cycle; wks= weeks; FU= follow up. For abbreviations on lab tests, see Section 9.4.1.

[a] Participants must be followed for at least 100 days after last dose of study treatment. Follow-up visit #1 (FU1) should occur 30 days from the last dose (+/- 7) days or can be performed on the date of discontinuation if that date is greater than 42 days from last dose. Follow-up visit #2 (FU2) occurs approximately 100 days (+/- 7 days) from last dose of study drug. Both Follow Up visits should be conducted in person.

[b] Survival Follow-up visits to occur every 3 months from Follow-up Visit #2. Survival visits may be conducted in person or by telephone. BMS may request that survival data be collected on all treated participants outside of the 3 month specified window. At the time of this request, each participant will be contacted to determine their survival status unless the participant has withdrawn consent for all contact.

3 INTRODUCTION

Renal cell carcinoma (RCC) is the eighth most common cancer in the world with an increasing incidence.[1] Globally, RCC occurs in more than 330,000 cases with approximately a third of the patients succumbing to their disease. More than 100,000 deaths occur annually, as a result of progression of metastatic disease. Despite the earlier detection of smaller kidney tumors, the rate of RCC-related mortality has increased suggesting that recurrence and advanced disease are responsible for mortality.[2,3] With the rise in RCC incidence, as well as mortality and morbidity associated with advanced RCC, medical need in this population remains a priority.

Over the last decade, an increased understanding of the biology of RCC has led to development of multiple agents that target specific growth pathways. The vascular endothelial growth factor (VEGF) pathway and targeted serine/threonine protein kinase therapies that block the mammalian target of rapamycin (mTOR) have been found to be important targets in RCC disease. Global health authorities (HAs) have approved multiple drugs targeting these pathways, including anti-VEGF agents, such as pazopanib, sorafenib, sunitinib, cabozantinib, and bevacizumab, and mTOR pathway inhibitors, such as temsirolimus and everolimus.[4] Additionally, recent innovation of treating cancer with immunotherapies has also expanded treatment options. Nivolumab, an anti-PD-1 antibody, given as monotherapy or in combination with the anti-CTLA-4 antibody, ipilimumab, has demonstrated clinical activity in multiple tumor types, including RCC.

To better understand treatment and patient outcomes, several academic groups have identified variables associated with survival and created prognostic models in mRCC. These risk models are commonly used for choosing therapies or selecting patients for treatment in clinical trials. The International Metastatic Renal Cell Carcinoma Database Consortium (IMDC) model stratifies patients into 3 prognostic groups, based on 6 adverse prognostic factors, into favorable (0 factors), intermediate (1-2 factors), and poor risk (3-6 factors) groups.[5] The currently available first-line agents for mRCC, which target the VEGF pathway, have shown limited efficacy in the intermediate and poor risk populations, yielding median overall survival of approximately 2 years or less.

Cabozantinib and nivolumab both share category 1 NCCN guideline recommendations (ie, uniform consensus that the treatment is appropriate, based on high-level evidence) for the treatment of previously treated mRCC patients.[6] Therefore, it is an appropriate next step to combine these agents and move them to the first-line setting in an attempt to improve clinical outcomes in patients with advanced RCC. This protocol CA2099ER will test the clinical activity of nivolumab combined with cabozantinib (doublet regimen). Given the different mechanisms of action of each of these agents, there is potential for distinct improvement in clinical efficacy.[7]

3.1 Study Rationale

Agents that target the VEGF pathway prevent tumor growth by inhibiting angiogenesis. A recent clinical biomarker study has shown that anti-VEGF therapy also affects the RCC tumor immune microenvironment. In this study, RCC tumor tissue from treatment-naive patients and those who had received prior bevacizumab or sunitinib were assessed for tumor immune cell infiltration.

Samples from patients with prior anti-VEGF therapy demonstrated increased infiltration of regulatory T cells (Tregs) and enhanced tumor PD-L1 expression, both of which were negatively associated with patient survival.[8] These effects suggest that the promotion of an immune suppressive tumor microenvironment may contribute to anti-VEGF therapy resistance and point to a rational strategy for combining anti-VEGF therapy with immunotherapies that target the PD-1/PD-L1 pathway. Indeed, the phase 1 study CA209016 demonstrated the synergistic activity of such combinations in mRCC. This study included a cohort of 26 participants with mRCC (19 of whom were previously untreated) who received sunitinib in combination with nivolumab 5 mg/kg Q3W. Using MSKCC risk criteria, 13 participants in this cohort had favorable risk disease, 12 had intermediate risk disease, and 1 had poor risk disease. After a minimum follow-up of 22 months, the objective response rate (ORR) was 42.3%, median progression-free survival (PFS) was 12.3 months, and median overall survival (OS) was 36.8 months in this cohort.[9] Similar efficacy results have been reported in other phase 1 studies which combine anti-VEGF agents with anti-PD-1 agents.[10,11]

As described in Section 3.2, cabozantinib is a novel tyrosine kinase inhibitor that, in addition to VEGFR, targets additional tyrosine kinases that are implicated in the biology of mRCC, such as c-MET and AXL. In a randomized phase 3 trial in patients with advanced RCC that had progressed after anti-VEGFR therapy, cabozantinib was shown to improve PFS and OS compared to everolimus, leading to its regulatory approval. Subsequently, a randomized phase 2 trial of cabozantinib vs sunitinib has demonstrated an improvement in ORR and PFS in intermediate- and poor-risk patients with previously untreated mRCC (see Section 3.2.1.4).

Cabozantinib has also been demonstrated to have effects on immune cells. In a study of 24 subjects with advanced urothelial carcinoma, cabozantinib treatment resulted in a decrease in circulating Tregs and increased PD-1 expression on Tregs. Low Tregs at baseline were also predictive of improved response to cabozantinib and survival.[12]

Given the promising clinical activity of cabozantinib in previously untreated mRCC and its potential immune effects, combining cabozantinib with nivolumab (in a doublet regimen) is a rational strategy to optimize first-line therapy in mRCC. An ongoing phase 1 study is evaluating both the doublet and triplet (nivolumab plus ipilimumab plus cabozantinib) regimens in patients with refractory advanced urothelial cancer or other genitourinary tumors, including mRCC, and has defined dosing for both regimens that produces acceptable safety and tolerability (Section 3.2.1.5). This 2-arm randomized phase 3 trial will determine if the combination doublet regimen (nivolumab combined with cabozantinib) produces greater clinical benefit than sunitinib, a standard of care agent for patients with previously untreated mRCC. In addition, this trial will reveal the adverse event profiles, quality of life measures, as well as exploratory biomarkers associated with these different first-line treatment regimens.

3.1.1 Research Hypothesis

Treatment with nivolumab combined with cabozantinib (doublet regimen) will demonstrate an improvement in PFS per BICR compared to sunitinib monotherapy in participants with previously untreated mRCC.

3.2 Background

Cancer immunotherapy rests on the premise that tumors can be recognized as foreign rather than as self and can be effectively attacked by an activated immune system. An effective immune response in this setting is thought to rely on immune surveillance of tumor antigens expressed on cancer cells that ultimately results in an adaptive immune response and cancer cell death. Meanwhile, tumor progression may depend upon acquisition of traits that allow cancer cells to evade immunosurveillance and escape effective innate and adaptive immune responses.[13,14,15] Current immunotherapy efforts attempt to break the apparent tolerance of the immune system to tumor cells and antigens by either introducing cancer antigens by therapeutic vaccination or by modulating regulatory checkpoints of the immune system. T-cell stimulation is a complex process involving the integration of numerous positive as well as negative co-stimulatory signals in addition to antigen recognition by the T-cell receptor (TCR).[16] Collectively, these signals govern the balance between T-cell activation and tolerance.

Programed death-1 (PD-1) is a member of the CD28 family of T-cell co-stimulatory receptors that also includes CD28, CTLA 4, ICOS, and BTLA.[17] PD-1 signaling has been shown to inhibit CD-28-mediated upregulation of IL-2, IL-10, IL-13, interferon-γ (IFN-γ) and Bcl-xL. PD-1 expression also been noted to inhibit T cell activation, and expansion of previously activated cells. Evidence for a negative regulatory role of PD-1 comes from studies of PD-1 deficient mice, which develop a variety of autoimmune phenotypes.[18] These results suggest that PD-1 blockade has the potential to activate anti-self T-cell responses, but these responses are variable and dependent upon various host genetic factors. Thus, PD-1 deficiency or inhibition is not accompanied by a universal loss of tolerance to self-antigens.

In vitro, nivolumab (BMS-936558) binds to PD-1 with high affinity (EC_{50} 0.39-2.62 nM), and inhibits the binding of PD-1 to its ligands PD-L1 and PD-L2 (IC_{50} ± 1 nM). Nivolumab binds specifically to PD-1 and not to related members of the CD28 family such as CD28, ICOS, CTLA-4 and BTLA. Blockade of the PD-1 pathway by nivolumab results in a reproducible enhancement of both proliferation and IFN-γ release in the mixed lymphocyte reaction (MLR). Using a cytomegalovirus (CMV) re stimulation assay with human peripheral blood mononuclear cells (PBMC), the effect of nivolumab on antigen specific recall response indicates that nivolumab augmented IFN-γ secretion from CMV specific memory T cells in a dose-dependent manner versus isotype-matched control. In vivo blockade of PD-1 by a murine analog of nivolumab enhances the anti-tumor immune response and result in tumor rejection in several immunocompetent mouse tumor models (MC38, SA1/N, and PAN02).[19]

CTLA-4, an activation-induced T-cell surface molecule, is a member of the CD28:B7 immunoglobulin superfamily that competes with CD28 for B7. CTLA-4 mediated signals are inhibitory and turn off T cell-dependent immune responses.[20] Ipilimumab is a fully human monoclonal IgG1κ that binds to the CTLA-4 antigen expressed on a subset of T cells from human and nonhuman primates. The proposed mechanism of action for ipilimumab is interference of the

interaction of CTLA-4 with B7 molecules on antigen presenting cells, with subsequent blockade of the inhibitory modulation of T-cell activation promoted by the CTLA 4/B7 interaction.

Cabozantinib inhibits multiple receptor tyrosine kinases (RTKs) implicated in tumor growth, metastasis, and angiogenesis.[21] The primary targets of cabozantinib implicated in mRCC are MET (c-MET), AXL and vascular endothelial growth factor receptor 2 (VEGFR2); additional targets identified in-vitro include RET, KIT, ROS1, TYRO3, MER, KIT, TRKB, FLT-3, and TIE-2. Both c-Met, AXL, and VEGFR2 are important mediators of tumor growth and tumor angiogenesis, and in vivo pharmacodynamic activity of cabozantinib against c-Met, and VEGFR2 has been demonstrated in both preclinical and clinical studies.[22,23,24,25]

In addition, preclinical and clinical observations have suggested that cabozantinib promotes an immunopermissive environment which might present an opportunity for synergistic effects from combination treatment with PD-1 checkpoint inhibitors. Specifically, treatment of tumor cells with cabozantinib in vitro led to increased tumor-cell expression of major histocompatibility complex (MHC) class 1 antigen and greater sensitivity of tumor cells to T-cell-mediated killing.[26] In a mouse tumor model, cabozantinib treatment led to increased peripheral CD8+ T-cell counts, decreased regulatory T-cells (Tregs) and myeloid-derived suppressor cells (MDSCs), and decreased Treg suppressor activity. Further, synergistic effects including increased CD8+ T-cell infiltration and decreased infiltration by MDSCs and tumor-assisted macrophages (TAMs) were observed when a poxviral-based cancer vaccine was administered in addition to cabozantinib in the mouse tumor model. Reductions in immunosuppressive Treg lymphocytes following treatment with cabozantinib have also been observed in subjects with advanced refractory urothelial cancer.[12] In a phase 2 study in metastatic triple-negative breast cancer, cabozantinib-treated subjects experienced a persistent increase in the fraction of circulating CD3+ T lymphocytes and a persistent decrease in the CD14+ monocytes possibly reflecting activation of systemic antitumor immunity.[27]

A detailed description of the chemistry, pharmacology, efficacy, and safety of nivolumab and ipilimumab in combination and cabozantinib is provided in the Nivolumab Investigator Brochure (IB) and Cabozantinib IB, respectively.[21,28,29]

3.2.1 Indication Background

3.2.1.1 Sunitinib in Renal Cell Carcinoma

Sunitinib is a VEGF receptor TKI that is approved and recommended for the treatment of mRCC across prognostic groups.[6,30] In a randomized phase 3 trial of sunitinib vs IFNα in treatment-naive subjects (including 36% with favorable risk, 57% with intermediate risk, and 7% with poor risk per MSKCC criteria), PFS (by independent radiology review) was significantly improved in the sunitinib group compared to the IFNα group (median PFS 11 vs 5 months, hazard ratio [HR] = 0.42; $p < 0.001$). ORR was also greater in the sunitinib group (31% vs 6%). Median OS was 26.4 months in the sunitinib group vs 21.8 months in the IFNα group (HR = 0.82; $p = 0.051$).[31] More recently, sunitinib was compared to pazopanib in treatment-naive subjects (including 27% favorable risk, 73% intermediate risk, and no poor risk) in the phase 3 COMPARZ study.[32] In this

non-inferiority study, sunitinib and pazopanib demonstrated similar median PFS (8.4 months for pazopanib vs 9.5 months for sunitinib, HR = 1.05) and median OS (28.4 months for pazopanib vs 29.3 months for sunitinib, HR = 0.91, p = 0.28). The ORR of pazopanib and sunitinib was 31% and 24%, respectively. The most common (≥ 20% frequency) adverse reactions include fatigue, asthenia, fever, diarrhea, nausea, mucositis/stomatitis, vomiting, dyspepsia, abdominal pain, constipation, hypertension, peripheral edema, rash, hand-foot syndrome, skin discoloration, dry skin, hair color changes, altered taste, headache, back pain, arthralgia, extremity pain, cough, dyspnea, anorexia, and bleeding.[33] Other important adverse reactions include hepatotoxicity, QT prolongation (including Torsades de Pointes), osteonecrosis of the jaw, tumor lysis syndrome, and thyroid dysfunction

3.2.1.2 Nivolumab Monotherapy in Renal Cell Carcinoma

Harnessing the immune system would be an attractive opportunity, together with efforts to find cell surface markers that can be used to trace and target dormant renal-cell carcinoma cells. Nivolumab monotherapy has been studied in participants with advanced RCC in several BMS-sponsored studies (phases 1 through 3): MDX1106-03, CA209009, CA209010, and CA209025. MDX1106-03 was a phase 1 refractory solid tumor trial, which included 34 participants with previously-treated advanced RCC who received nivolumab at 1 mg/kg (n = 18) or 10 mg/kg (n = 16) given every 2 weeks.[24,34,35] Median progression-free survival (PFS) was 7.3 months. In both the 1 mg/kg and 10 mg/kg cohorts, approximately 30% of participants experienced an objective response with median duration of response of 12.9 months. Responses were generally rapid with a median time to response of 16 weeks. Notably, responses could occur after treatment cessation and persist off treatment. Median overall survival (mOS) was 22.4 months. These results were promising given that many of the participants were heavily pre-treated with 71% having had 2 or more lines of therapy. Treatment-related adverse events (AEs) of any grade were observed in 85% of RCC participants, the most common being fatigue (41%), rash (27%), diarrhea (18%), and pruritus (18%). Grade 3-4 treatment-related AEs were observed in 18% of RCC participants. The spectrum, frequency, and severity of treatment-related AEs were similar in the RCC population compared to the overall study population and were similar across dose levels.

In order to identify a potential dose-response relationship in RCC, a randomized phase 2 study (CA209010) was conducted in 168 participants with advanced RCC previously treated with an antiangiogenic therapy who received nivolumab at 0.3 mg (n = 60), 2 mg/kg (n = 24) or 10 mg/kg (n=54) given every 3 weeks.[36] No dose response relationship was found as measured by PFS, with median PFS of 2.7, 4.0, and 4.2 months for the 0.3, 2, and 10 mg/kg groups, respectively (P = 0.9). ORR was 20%, 22%, and 20% in the 0.3, 2, and 10 mg/kg groups, respectively. Median time to achievement of an objective response was 2.8-3.0 months. The median duration of response was 22.3 months (4.8, NR) in the 10 mg/kg arm and not yet reached in the 2 lower dose cohorts. Median OS was at 18.2 to 25.5 months, with a minimum of follow-up of 24 months. Fatigue was the most frequent toxicity (22-35%). No new toxicities were identified with 11% experiencing Grade 3-4 treatment-related AEs, none of which were due to pneumonitis. Treatment-related AEs led to discontinuation of study drug in 7% of participants.

A parallel biomarker-focused trial, CA209009, using the same 3 nivolumab dose levels (03, 2, and 10 mg/kg every 3 weeks) was executed to explore predictors of response and identify mechanisms of resistance.[37] This study included 67 participants with previously-treated, advanced RCC who were randomized to one of the 3 nivolumab dose groups and 24 participants with previously-untreated RCC who received nivolumab at 10 mg/kg every 3 weeks. The results mirrored the efficacy and toxicity profile of CA209010 with an overall ORR of 18% in previously-treated participants, and 13% in previously untreated participants and disease stabilization in another 32% of previously treated and untreated participants. At 24 weeks, 36% of participants were free from progression. Of 56 participants with evaluable pretreatment tumor samples, 18 (32%) had ≥ 5% PD-L1 tumor expression. ORR was 22% among those with ≥ 5% PD-L1 tumor expression versus 8% among those with < 5% PD-L1 tumor expression.

Based on the clinical activity of nivolumab observed in these phase 1 and 2 studies, a large phase 3 trial (CA209025) was conducted in 821 participants with advanced RCC previously treated with 1 or 2 antiangiogenic therapies who were randomized to receive nivolumab 3 mg/kg every 2 weeks or everolimus 10 mg daily. A planned interim analysis, after a minimum of follow-up of 14 months, demonstrated a statistically significant and clinically meaningful improvement in OS of nivolumab monotherapy vs everolimus (median OS, 25.0 months vs 19.6 months, respectively; HR 0.73 [98.5% CI, 0.57 to 0.93, P = 0.002). ORR was 25% for nivolumab vs 5% for everolimus. Additional efficacy results are presented in Table 3.2.1.2-1. Among 756 participants with quantifiable PD-L1 tumor expression in pretreatment samples, 24% had ≥1% PD-L1 expression. Among participants with ≥1% PD-L1 expression, median OS was 21.8 months in the nivolumab group and 18.8 months in the everolimus group (HR, 0.79; 95% CI, 0.53 to 1.17). Among participants with <1% PD-L1 expression, the median OS was 27.4 months in the nivolumab group and 21.2 months in the everolimus group (HR, 0.77; 95% CI 0.60 to 0.97). No new safety concerns were identified, and nivolumab monotherapy showed a favorable safety profile as compared to everolimus, evidenced by the lower rates of drug-related AEs (all grades, 79% vs 88%; Grade 3-4, 19%-37%, respectively) and drug-related AEs leading to discontinuation (all grades, 8% vs 13%, respectively) in the nivolumab group. These results were the basis for regulatory approval of nivolumab monotherapy in advanced RCC.

Table 3.2.1.2-1: Summary of Efficacy Results - All Randomized Subjects - CA209025

Efficacy Parameters	Nivolumab (N = 410)	Everolimus (N = 411)
Primary Endpoint		
Overall Survival		
Events , n (%)	183/410 (44.6)	215/411 (52.3)
Stratified log-rank test P value[a,b]	0.0018	
HR (98.52% CI)[c]	0.73 (0.57, 0.93)	
Median (95% CI), months[d]	25.00 (21.75, NR)	19.55 (17.64, 23.06)

Table 3.2.1.2-1: Summary of Efficacy Results - All Randomized Subjects - CA209025

Efficacy Parameters	Nivolumab (N = 410)	Everolimus (N = 411)
Rate at 6 months (95% CI), %[d]	89.2 (85.7, 91.8)	81.2 (77.0, 84.7)
Rate at 12 months (95% CI), %[d]	76.0 (71.5, 79.9)	66.7 (61.8, 71.0)
Secondary Endpoints		
Objective Response Rate per Investigator (CR + PR)[e]		
N (%)	103 (25.1)	22 (5.4)
95% CI[f]	(21.0, 29.6)	(3.4, 8.0)
Odds ratio estimate (95%CI)[g,h]	5.98 (3.68, 9.72)	
P Value[i]	< 0.0001	
Duration of response[e]		
Ongoing responders, n/N (%)	49/103 (47.6)	10/22 (45.5)
Median (95% CI), months[d]	11.99 (7.85, 23.03)	11.99 (6.44, NR)
Min, Max[j]	0.0, 27.6+	0.0+, 22.2+
Progression-free survival		
Events, n (%)	318 (77.6)	322 (78.3)
Stratified log-rank test p value[a]	0.1135	
HR (95% CI)[c]	0.88 (0.75, 1.03)	
Median 95% CI)	4.60 (3.71, 5.39)	4.44 (3.71, 5.52)

[a] Log-rank test stratified by the MSKCC risk group (poor vs intermediate vs favorable), the number of prior antiangiogenic therapies in the advanced/metastatic setting (1 vs 2), and the region (W. Europe, US/Canada vs Rest of the World) as entered into the IVRS.

[b] Based on the 398 observed deaths and O'Brien-Fleming alpha spending function, the boundary for statistical significance requires the P value to be less than 0.0148.

[c] Stratified Cox proportional hazard model. Hazard ratio is nivolumab over everolimus.

[d] Based on Kaplan-Meier Estimates.

[e] The confirmed ORR was 88/410 (21.5%) in the nivolumab group and 16/411 (3.9%) in the everolimus group (stratified CMH test P value < 0.0001), with a median DOR of 23.03 months in the nivolumab group and 13.73 months in the everolimus group.

[f] CR+PR, confidence interval based on the Clopper and Pearson method.

[g] Cochran-Mantel-Haenzel test stratified by the MSKCC risk group (poor vs intermediate vs favorable), the number of prior anti-angiogenic therapies in the advanced/metastatic setting (X vs 2) and the region (Western Europe vs US/Canada vs Rest of the World) as entered into the IVRS.

[h] Ratio of nivolumab over everolimus

[i] Two-sided p value from CMH test for the comparison of the odds ratio of nivolumab over everolimus.

[j] Symbol + indicated a censored value.

3.2.1.3 Nivolumab Plus Ipilimumab in Renal Cell Carcinoma

Promising safety and efficacy results were also observed with the combination of nivolumab and ipilimumab in the advanced RCC population in study CA209016,[9] a phase 1 dose-escalation study of nivolumab in combination with VEGFR-TKIs or ipilimumab in participants with metastatic RCC. Treatment-experienced and -naive participants with metastatic RCC were randomized to receive nivolumab 3 mg/kg + ipilimumab 1 mg/kg (arm N3 + I1) or nivolumab 1 mg/kg + ipilimumab 3 mg/kg (arm N1 + I3) IV Q3W for 4 doses then nivolumab 3 mg/kg IV Q2W until progression/toxicity. In Arm N1+I3, 25 out of 47 participants (53%) were treatment-naive, and 61.7% were intermediate risk according to MSKCC criteria, and 4.3% were categorized as poor risk. In Arm N1 + I3, 21 of 47 participants (45%) were treatment-naive, and 66.0% were in the intermediate risk category, and 6.4% were categorized as poor risk. The primary objective was to assess safety/tolerability; secondary objective was to assess antitumor activity.

After a minimum of 22 months, the level of clinical activity, as measured by confirmed ORR, for the combination of nivolumab and ipilimumab in CA209016 was substantially greater than that observed in studies of either nivolumab monotherapy (Section 3.2.1.2) in metastatic RCC, including in the treatment-naive subpopulation. The dosing regimen including nivolumab 3 mg/kg combined with ipilimumab 1 mg/kg (N3 + I1) was chosen for further clinical evaluation because it exhibited similar clinical activity to nivolumab 1 mg/kg combined with ipilimumab 3 mg/kg (N1 + I3) but had a more favorable safety profile.

Table 3.2.1.3-1: Antitumor Activity in All Participants (CA209016)[9]

	N3 + I1 (n = 22) Previously Treated	N3 + I1 (n = 25) Treatment-Naive	N1 + I3 (n = 26) Previously Treated	N1 + I 3 (n = 21) Treatment-naive
Confirmed ORR, n (%) (95% CI)	10 (45.5)	9 (36.0) 18.0, 57.5)	10 (38.5)	9 (42.9) (21.8, 66.0)
Median duration of response, weeks (range)	60.1 (9.29, NA)	88.7 (30.00, 105.00)	74.4 (12.29, 108.29)	5 (55.6)
Ongoing responses, % (n/N)	4 (40.0)	4 (44.4)	2 (20.0)	NR (23.57, NA)
Best objective response, n (%)				
Complete response	3 (13.6)	2 (8.0)	0	0
Partial response	7 (31.8)	7 (28.0)	10 (38.5)	9 (42.9)
Stable disease	6 (27.3)	13 (52.0)	11 (42.3)	6 (28.8)
Progressive disease	6 (27.3)	6 (27.3)	3 (11.5)	5 (23.8)
Unable to determine	0	0	2 (7.7)	1 (4.8)
PFS Median months (CI)	6.6 (1.41, 16.39)	8.3 (3.55, 19.29)	10.1 (5.42, 20.76)	8.5 (2.00, NA)
6-month PFS Median months (CI)	54.5 (32.1, 72.4)	56.5 (34.3, 73.8)	65.4 (44.0, 80.3)	61.9 (38.1, 78.8)

Table 3.2.1.3-1: Antitumor Activity in All Participants (CA209016)[9]

	N3 + I1 (n = 22) Previously Treated	N3 + I1 (n = 25) Treatment-Naive	N1 + I3 (n = 26) Previously Treated	N1 + I 3 (n = 21) Treatment-naive
Median OS	NR (10.02, NA)	NR (26.68, NA)	30.9 (25.99, NA)	NR (17.45, NA)

Abbreviations: NR = Not Reached

Among the 91 participants treated with nivolumab + ipilimumab combination in CA209016 who provided evaluable baseline tumor samples, 37.4% had ≥ 1% PD-L1 tumor expression, and 16.5% had ≥5% PD-L1 tumor expression. ORR was 47.1% among participants with ≥1% PD-L1 expression and 36.8% among participants with < 1% PD-L1 expression. Among participants with ≥5% PD-L1 expression, ORR was 40.0%.[9]

Among all treated patients in the Arms N3 + I1 and N1 + N3, AEs were seen in 43/47 (91.5%) participants in the N3 + I1 arm and 45/47 (95.7%) participants in the N1 + I3 arm. In the N1 + I3 arm, the most frequently reported drug-related AEs were fatigue (51.1%); rash, and pruritus (each 31.9%); nausea (27.7%); arthralgia (25.5%). In the N1 + I3 arm, the most frequently reported drug related AEs were fatigue (68.1%); diarrhea, and nausea (each 44.7%); pruritus (36.2%); lipase increased (34%); AST increased (31.9%); ALT increased, and decreased appetite (29.8%); hypothyroidism (27.7%); and rash (25.5%). In the N3 + I1 arm, the most frequently reported, Grade 3-4 drug-related AE was lipase increased (14.9%). In the N1 + I3 arm, the most frequently reported Grade 3-4 drug related AEs were lipase increased (27.7%); ALT increased (21.3%); diarrhea, and colitis (14.9%); AST increased (12.8%).[9]

Treatment-related AEs (including Grade 3-4), treatment-related AEs leading to discontinuation, and treatment-related SAEs all occurred more commonly in participants in the N1 + I3 arm than in the N3 + I1 arm (Table 3.2.1.3-2).[28]

Table 3.2.1.3-2: Summary of Safety Results - All Treated Subjects

	Arm I-1 IPI1 + NIV3 N = 47		Arm I-3 IP13 + NIV1 N = 47	
Death, n (%)	16 (34.0)		18 (38.3)	
Within 30 Days of Last Dose	0		1 (2.1)	
Within 100 Days of Last Dose	3 (6.4)		4 (8.5)	
Due to Study Drug Toxicity	0		0	
	Any Grade	**Grade 3-4**	**Any Grade**	**Grade 3-4**
All-causality SAEs, n (%)	29 (61.7)	20 (42.6)	30 (63.8)	24 (51.0)
Drug-related SAEs, n (%)	11 (23.4)	9 (19.1)	16 (34.0)	16 (34.0)

Table 3.2.1.3-2: Summary of Safety Results - All Treated Subjects

	Arm I-1 IPI1 + NIV3 N = 47		Arm I-3 IP13 + NIV1 N = 47	
All-causality AEs Leading to Discontinuation, n (%)	5 (10.6)	3 (6.4)	15 (31.9)	11 (23.4)
Drug-related AEs Leading to Discontinuation, n (%)	5 (10.6)	3 (6.4)	13 (27.7)	9 (19.1)
All-causality AEs, n (%)	47 (100.0)	33 (70.2)	47 (100.0)	34 (72.3)
Drug-related AEs, n (%)	43 (91.5)	18 (38.3)	45 (95.7)	29 (61.7)
All-causality Select AEs, within 30 Days of Last Dose, by Category, n (%)	**Any Grade**	**Grade 3-4**	**Any Grade**	**Grade 3-4**
Endocrine	14 (29.8)	3 (6.4)	19 (40.4)	0
Gastrointestinal	16 (34.0)	3 (6.4)	25 (53.2)	12 (25.5)
Hepatic	11 (23.4)	3 (6.4)	15 (31.9)	8 (17.0)
Pulmonary	3 (6.4)	0	5 (10.6)	0
Renal	11 (23.4)	2 (4.3)	10 (21.3)	2 (4.3)
Skin	29 (61.7)	1 (2.1)	33 (70.2)	1 (2.1)
Hypersensitivity/Infusion Reactions	5 (10.6)	0	3 (6.4)	0
Drug-related Select AEs, within 30 Days of Last Dose, by Category, n (%)	**Any Grade**	**Grade 3-4**	**Any Grade**	**Grade 3-4**
Endocrine	13 (27.7)	2 (4.3)	19 (40.4)	0
Gastrointestinal	12 (25.5)	2 (4.3)	21 (44.7)	11 (23.4)
Hepatic	9 (19.1)	3 (6.4)	13 (27.7)	8 (17.0)
Pulmonary	3 (6.4)	0	5 (10.6)	0
Renal	9 (19.1)	2 (4.3)	6 (12.8)	1 (2.1)
Skin	23 (48.9)	0	28 (59.6)	1 (2.1)
Hypersensitivity/Infusion Reactions	5 (10.6)	0	3 (6.4)	0
All-causality Immune-mediated AEs, by Category ***Immune-mediated AEs Treated with Immune-modulating medication***	**Any Grade**	**Grade 3-4**	**Any Grade**	**Grade 3-4**
Diarrhea/Colitis	3 (6.4)	2 (4.3)	12 (25.5)	10 (21.3)
Hepatitis	5 (10.6)	2 (4.3)	11 (23.4)	8 (17.0)
Pneumonitis	1 (2.1)	0	5 (10.6)	0
Nephritis and Renal Dysfunction	2 (4.3)	1 (2.1)	1 (2.1)	0

Table 3.2.1.3-2: Summary of Safety Results - All Treated Subjects

	Arm I-1 IPI1 + NIV3 N = 47		Arm I-3 IP13 + NIV1 N = 47	
Rash	8 (17.0)	1 (2.1)	9 (19.1)	1 (2.1)
Hypersensitivity	0	0	0	0
Immune-Mediated Endocrine AEs Treated with or without Immune-Modulating Medications	**Any Grade**	**Grade 3-4**	**Any Grade**	**Grade 3-4**
Adrenal Insufficiency	3 (6.4)	1 (2.1)	6 (12.8)	0
Hypophysitis	1 (2.1)	1 (2.1)	2 (4.3)	0
Hypothyroidism/Thyroiditis	10 (21.3)	0	14 (29.8)	0
Hyperthyroidism	4 (8.5)	1 (2.1)	8 (17.0)	0
Diabetes Mellitus	0	0	0	0

MedDRA version 18.1; CTC version 4.0. All events are within 100 days of the last dose of study drug, unless otherwise indicated. Sources: Table S.6.2A (deaths), Table S.6.3A (all-causality SAEs), Table S.6.3B (drug-related SAEs), Table S.6.4B (all-causality AEs leading to discontinuation), Table S.6.4D (drug-related AEs leading to discontinuation), Table S.6.2 (all-causality AEs), Table S.6.3.1 (drug-related AEs), Table S.6.101 (all-causality select AEs), Table S.6.105 (all-causality endocrine select AEs), Table S.6.103 (drug-related select AEs), Table S.6.107 (drug-related endocrine select AEs), Table S.6.202 (all-causality IMAEs with exception of endocrine), and Table S.6.204 (all-causality endocrine IMAEs)

A large phase 3 trial, study CA209214, is currently ongoing to determine if nivolumab combined with ipilimumab (N3+ I1 regimen) improves PFS, OS, and ORR vs sunitinib, in participants with previously untreated, advanced or metastatic RCC.

A detailed description of the chemistry, pharmacology, efficacy, and safety of nivolumab and ipilimumab in combination is provided in the Nivolumab Investigator Brochure and Product Label.[28] Note that the nivolumab and ipilimumab combination is currently approved for unresectable or metastatic melanoma, using the N1 + I3 regimen.

3.2.1.4 Cabozantinib Monotherapy in Renal Cell Carcinoma

Cabozantinib is a small molecule inhibitor of the tyrosine kinases c-Met, AXL, and VEGFR2, and has been shown to reduce tumor growth, metastasis, and angiogenesis. Cabozantinib has been evaluated in both first-line and second-line settings in advanced and metastatic RCC.

In participants with advanced renal cell carcinoma (RCC) who progressed after previous VEGFR tyrosine-kinase inhibitor (VEGFR-TKI) treatment, the randomized phase 3 METEOR trial compared the efficacy and safety of cabozantinib monotherapy at a daily dose of 60 mg versus the mTOR inhibitor everolimus at a daily dose of 10 mg.[38,39] In the trial, 658 participants were randomized to receive cabozantinib (n = 330) or everolimus (n = 328). After a minimum follow-up of 11 months in the first 375 randomized subjects, the primary endpoint of PFS (by independent radiology review) was 7.4 months in the cabozantinib arm vs 3.8 months in the everolimus arm (HR 0.58 [95% CI 0.45 to 0.75, p<0.001). Using Memorial Sloan Kettering Cancer Center (MSKCC) criteria, subgroup analyses of risk groups demonstrated PFS benefit in participants with

favorable risk (HR 0.54 [95% CI 0.37, 0.79]), intermediate risk (HR 0.56 [95% CI 0.37, 0.84]), and poor risk (HR 0.84 [95% CI 0.46, 1.53]). ORR was 17% in the cabozantinib arm vs 5% in the everolimus arm ($p<0.001$). An interim OS analysis at the time of this final PFS analysis demonstrated a trend toward longer OS (HR 0.67, $p=0.005$, with $p\leq0.0019$ required for statistical significance). A subsequent OS analysis, performed after a median follow-up of approximately 19 months in all 658 randomized subjects, demonstrated a significant improvement in OS, with median OS of 21.4 months in the cabozantinib arm and 16.5 months in the everolimus arm (HR 0.66 [95% CI 0.53–0.83]; $p = 0.00026$). Subgroup analyses of OS according to MSKCC risk group were consistent with the results for the overall population. The most common grade 3 or 4 adverse events included hypertension (49 [15%] in the cabozantinib group vs 12 [4%] in the everolimus group), diarrhea (43 [13%] vs 7 [2%]), fatigue (36 [11%] vs 24 [7%]), palmar-plantar erythrodysesthesia syndrome (27 [8%] vs 3 [1%]), anemia (19 [6%] vs 53 [17%]), hyperglycemia (3 [1%] vs 16 [5%]), and hypomagnesemia (16 [5%] vs none). Serious adverse events grade ≥ 3 occurred in 130 (39%) participants in the cabozantinib group and in 129 (40%) in the everolimus group. One treatment-related death occurred in the cabozantinib group (death; not otherwise specified) and two occurred in the everolimus group (1 aspergillus infection and 1 pneumonia aspiration).[38,39]

In treatment naive participants, cabozantinib was also evaluated in a randomized phase 2 multicenter trial against sunitinib as first-line therapy in participants with advanced or metastatic RCC.[40] Participants were required to have either intermediate or poor risk disease according to IMDC criteria and were randomized in 1:1 ratio to cabozantinib (n = 79) or sunitinib (n = 78). Investigator-assessed PFS was the primary endpoint. Compared with sunitinib, cabozantinib treatment significantly increased median PFS (8.2 v 5.6 months) and was associated with a 34% reduction in rate of progression or death (adjusted hazard ratio, 0.66; 95% CI, 0.46 to 0.95; one-sided P = .012). ORR was 46% (95% CI, 34 to 57) for cabozantinib vs 18% (95% CI, 10 to 28) for sunitinib. Median OS was 30.3 months for cabozantinib vs 21.8 months for sunitinib (adjusted HR 0.80, 95% CI 0.50, 1.26). All-causality grade 3 or 4 adverse events were 67% for cabozantinib and 68% for sunitinib and included diarrhea (cabozantinib, 10% v sunitinib, 11%), fatigue (6% v 15%), hypertension (28% v 22%), palmar-plantar erythrodysesthesia (8% v 4%), and hematologic adverse events (3% v 22%). Treatment-related Grade 5 events occurred in 3 participants in the cabozantinib arm (acute kidney injury, sepsis, and jejunal perforation) and 3 participants in the sunitinib arm (sepsis, respiratory failure, and vascular disorders).

3.2.1.5 Nivolumab Combined with Cabozantinib or Nivolumab and Ipilimumab Combined with Cabozantinib

CA2099ER Global Revised Protocol 01

Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B. Participants previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinical planned events, per protocol.

The ongoing phase 1 trial (NCT02496208) evaluating the safety and efficacy of nivolumab combined with cabozantinib (doublet regimen) and nivolumab and ipilimumab combined with

cabozantinib (triplet regimen) in participants with refractory metastatic urothelial carcinoma (mUC) and other genitourinary tumors has reported interim results.[41] The primary objective was to determine the dose limiting toxicity and recommended phase 2 dose of the doublet and triplet regimens.

At the time of an update in February 2017, 48 participants (enrolled from 22-Jul-2015 to 31-Dec-2016) had been treated, including 19 with urothelial carcinoma; 9 with castration-resistance prostate cancer; 4 with urachal adenocarcinoma; 4 with germ cell tumor; 4 with penile cancer; 2 with squamous cell carcinoma of the bladder/urethra; 2 with clear cell RCC; 2 with sarcomatoid RCC; 1 with trophoblastic tumor; and 1 with Sertoli cell tumor.[41] Part 1 of the study was the dose escalation for the doublet regimen and enrolled first, followed by Part 2, which was the dose escalation for the triplet regimen (which dosed nivolumab and ipilimumab Q3W for the first 4 doses, followed by nivolumab alone Q2W). Thirty participants were treated with the doublet, and 18 were treated with triplet. Both the doublet and triplet regimens were found to be safe and tolerable. The recommended phase 2 dose for the doublet was cabozantinib 40 mg/day + nivolumab 3 mg/kg. The recommended phase 2 dose for the triplet was cabozantinib 40 mg/day + nivolumab 3 mg/kg + ipilimumab 1 mg/kg.

In the doublet arm, the most common treatment-related AEs of any grade were ALT increased (67%); fatigue (63%); diarrhea (60%); hypothyroidism (57%); ALT increased (50%); and anorexia (47%). The most common treatment-related Grade 3-4 AEs were neutropenia (17%); hypophosphatemia and lipase increase (13% each); hypertension (10%); fatigue, diarrhea, thrombocytopenia, and dehydration (7% each). One participant developed immune-mediated Grade 3 aseptic meningitis.

In the triplet arm, the most common treatment-related AEs of any grade were fatigue (72%); diarrhea and anorexia (61% each); ALT increased, hypophosphatemia, dysgeusia, and lipase increased (44% each). The most common treatment-related Grade 3-4 AEs were hypertension and hypophosphatemia (17% each); and fatigue, hyponatremia, nausea, and lipase increase (13% each). One participant developed immune-mediated Grade 3 colitis.

Among 43 participants who were evaluable for response, the ORR was 30%. The ORR was 38% among 26 evaluable participants treated with the doublet and 18% among 17 participants treated with the triplet. One of 2 participants with sarcomatoid RCC achieved a response. Neither of the 2 participants with clear cell RCC were evaluable for response at the time of the analysis.

The doublet and triplet expansion cohorts are continuing to enroll participants with advanced urothelial carcinoma and RCC. Updated safety and efficacy results after additional enrollment and longer follow-up are awaited.

3.3 Benefit/Risk Assessment

The currently available first-line agents for treatment of mRCC are associated with median OS of 43.2 months in patients with favorable risk disease (which accounts for approximate 25% of all untreated mRCC), 22.5 months in those with intermediate risk disease and only 7.8 months in those with poor risk disease according to IMDC criteria.[42] Recently, cabozantinib demonstrated a

significant improvement in investigator-assessed PFS and ORR over standard-of-care sunitinib as first-line therapy in a randomized phase 2 study in participants with intermediate- or poor-risk mRCC. Nivolumab plus ipilimumab showed a significant improvement in OS in participants with intermediate and poor risk mRCC, compared with sunitinib. At this moment, sunitinib is still one of the treatment options for first-line mRCC in all risk groups. Further improvement of efficacy and safety in all risk groups using a combination of agents with synergistic mechanisms of action is highly desirable. [43]

Based on their different mechanisms of action and the potential immune effects associated with anti-VEGF treatment, combining nivolumab with cabozantinib (doublet regimen) may produce synergistic clinical activity and provide improved benefit over standard of care sunitinib monotherapy in mRCC. As mentioned in Section 3.1, several phase 1 studies combining anti-PD-1 agents with anti-VEGF agents have demonstrated greater clinical activity than has been observed historically in trials of these agents given as monotherapy. Although additional risk may be involved with combination therapies, early safety results from the ongoing phase 1 combination study in refractory genitourinary cancers evaluating the safety and efficacy of nivolumab combined with cabozantinib (doublet regimen) suggest that safety profiles appear acceptable. Therefore, overall benefit/risk is acceptable for doublet combination compared to sunitinib in participants with mRCC in all risk groups.

Although the relative benefit of cabozantinib vs sunitinib in participants with previously untreated, favorable risk mRCC was not evaluated in the CABOSUN study, subgroup analyses from the METEOR study suggest that the PFS and OS benefit of cabozantinib in favorable risk disease is similar to that in intermediate risk disease (see Section 3.2.1.3). Based on the preliminary efficacy results from the phase 1 combination study, which included participants with refractory genitourinary tumors, nivolumab combined with cabozantinib (doublet) is anticipated to produce clinical activity that is greater than that observed with sunitinib monotherapy, even in favorable risk disease. Therefore, this study will also include participants with favorable risk disease, to comprise approximately 25% of the overall population.

Overall, the safety profile of nivolumab monotherapy is manageable and generally consistent across completed and ongoing clinical trials with no MTD reached at any dose tested up to 10 mg/kg. Most AEs were low-grade (Grade 1 to 2) with relatively few related high-grade (Grade 3 to 4) AEs. There was no pattern in the incidence, severity, or causality of AEs with respect to nivolumab dose level.

A pattern of immune-related adverse events has been defined, for which management algorithms have been developed; these are provided in Appendix 5 and within the Investigator Brochure.[28] Most high-grade events were manageable with the use of corticosteroids or hormone replacement therapy (endocrinopathies) as instructed in these algorithms.

More detailed information about the known and expected benefits and risks and reasonably anticipated adverse events (AEs) of nivolumab and cabozantinib may be found in their respective Investigator Brochures.[21,28]

4 OBJECTIVES AND ENDPOINTS

Table 4-1: Objectives and Endpoints

Objective	Endpoint
Primary	
To compare progression-free survival (PFS) per BICR of nivolumab combined with cabozantinib (Arm A: doublet) with sunitinib (Arm C) in all randomized participants.	The primary endpoint of this study is to compare PFS per BICR of Arm A versus Arm C in all randomized participants. PFS is defined as the time between the date of randomization and the first date of the documented progression, or death due to any cause, whichever occurs first. Participants who die without a reported progression (and die without start of subsequent anti-cancer therapy) will be considered to have progressed on the date of their death. Participants who did not progress or die will be censored on the date of their last evaluable tumor assessment on or prior to initiation of subsequent anti-cancer therapy. Participants who did not have any on study tumor assessments and did not die will be censored on their date of randomization. Participants who started anti-cancer therapy without a prior reported progression will be censored on the date of their last evaluable tumor assessment on or prior to the initiation of first subsequent anti-cancer therapy.
Secondary	
To compare overall survival (OS) of Arm A with Arm C in all randomized participants.	The first secondary endpoint is to compare OS of Arm A versus Arm C in all randomized participants. OS is defined as the time between the date of randomization and the date of death due to any cause. A participant who has not died will be censored at the last known alive date.
To evaluate the objective response rate (ORR) per BICR in all randomized participants.	The second secondary endpoint is to describe ORR per BICR in all randomized participants. ORR is defined as the proportion of randomized participants who achieve a best response of complete response (CR) or partial response (PR) using the RECIST 1.1 criteria. Best overall response (BOR) is defined as the best response designation recorded between the date of randomization and the date of objectively documented progression per RECIST 1.1 or the date of subsequent therapy (including tumor-directed radiotherapy and tumor-directed surgery), whichever occurs first. For participants without document progression or subsequent therapy, all available response designations will contribute to the BOR assessment. Duration of response (DOR) is defined as the time between the date of first confirmed documented response (CR or PR) to the date of first documented tumor progression (per RECIST 1.1) or death due to any cause, whichever occurs first. Participants who neither progress nor die will be censored on the date of their last tumor assessment. Responders who started anti-cancer therapy without a prior reported progression will be censored on the date of their last evaluable tumor assessment prior to the initiation of first subsequent anti-cancer therapy. Time to response (TTR) is defined as the time from randomization to the date of the first confirmed documented response (CR or PR), as assessed by BICR. DOR and TTR will be evaluated for responders (CR or PR) only.

Table 4-1: Objectives and Endpoints

Objective	Endpoint
To assess overall safety and tolerability in all treated participants.	As measured by the incidence of adverse events (AEs), serious adverse events (SAEs), AEs leading to discontinuation, deaths, laboratory abnormalities and changes from baseline.
Exploratory	
To explore potential predictive biomarkers of clinical response to nivolumab and cabozantinib combination.	Analysis of tumor specimens and blood samples for proteins and genes involved in regulating immune response (eg, PD-1, PD-L1, PD-L2, CXCL10, MET). Other exploratory endpoints for biomarkers, pharmacogenomics, and immunogenicity are described in Section 9.8.
To evaluate health related quality of life (HRQoL).	Assessed by the NCCN Functional Assessment of Cancer Therapy- Kidney Symptom Index (FKSI-19) and the EuroQoL Group's EQ-5D (3L version) is described in Section 9.1.4.
To characterize the pharmacokinetics of nivolumab and cabozantinib and explore exposure response relationships, if applicable.	Population PK parameters, exposure-response relationship between select PK measures of exposure and safety and efficacy endpoints, if applicable
To characterize the immunogenicity of nivolumab	Incidence of anti-nivolumab -antibodies and their potential relationship with safety and efficacy endpoints
To assess PFS after next line of treatment (PFS-2) in each arm	PFS-2 is defined as the time from randomization to the date of investigator-defined documented second objective disease progression after second-line therapy or death due to any cause, whichever comes first. Clinical deterioration will not be considered as progression. A subject who neither progresses nor dies will be censored on the date of his/her last adequate tumor assessment or last follow-up for progression/subsequent therapy. A subject who does not have any post-baseline tumor assessments and who has not died will be censored on the date at which he/she was randomized.

5 STUDY DESIGN

5.1 Overall Design

CA2099ER Global Revised Protocol 01

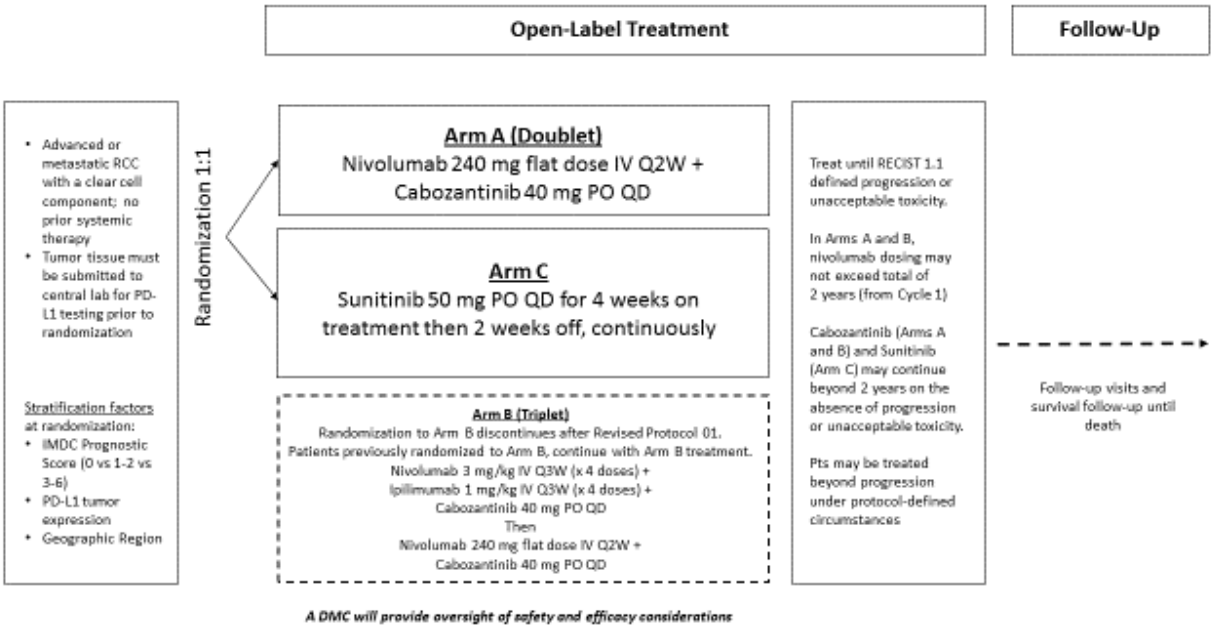
Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B. Participants previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinical planned events, per protocol.

This is an open label, randomized trial of nivolumab combined with cabozantinib (doublet regimen) versus sunitinib in participants with previously untreated (first line) advanced or metastatic RCC. Participants will be randomized between Arm A and Arm C in a 1:1 ratio with approximately 638 participants (319 per arm) capped at approximate 25% to represent the normal frequency of favorable risk group in mRCC. The rest of the randomized participants will provide approximately 478 intermediate/poor risk randomized participants (239 per arm). Participants will be stratified at the time of randomization by IMDC prognostic score (0 [favorable risk] versus-1-

2 [intermediate risk] versus 3-6 [poor risk]), PD-L1 tumor expression (≥ 1% versus < 1% or indeterminate), and region (US/Canada/Western Europe/Northern Europe versus rest of the world [ROW]).

The study design schematic is presented in Figure 5.1-1.

Figure 5.1-1: Study Design Schematic



Abbreviations: DMC= data monitoring committee; IMDC= International Metastatic Renal Cell Carcinoma Database Consortium; IV=intravenous; PD-L1= programmed death-ligand 1; PO= orally by mouth; Pts=patients/participants; Q2W=every 2 weeks; Q3W=every 3 weeks; QD= once daily; RCC=renal cell carcinoma.

Arm B is shown in the design schematic because after implementation of Revised Protocol 01, there will be be patients on treatment from the original protocol version 08 Mar 2017.

This study will consist of 3 stages: screening, treatment, and follow up phase.

Screening stage: Screening begins by establishing the participant's initial eligibility and signing of the informed consent (ICF). Sufficient, recent tumor tissue, preferably obtained within 3 months but no more than 12 months prior to enrollment, from a metastatic tumor lesion or from a primary tumor lesion which has not been previously irradiated (formalin-fixed paraffin-embedded block or 20 unstained slides: a minimum of 10 slides, obtained from core biopsy, punch biopsy, excisional biopsy or surgical specimen) will be submitted to the central laboratory. Upon receipt of the tumor sample, the central lab will determine PD-L1 expression level by immunohistochemistry (IHC) testing (see Section 9.8.2). In order to be randomized in the Interactive Response Technology (IRT) system, a participant must be classified as PD-L1 expression ≥ 1%, PD-L1 expression < 1%,

or PD-L1 expression indeterminate. Sites will be informed when the submitted tumor sample is insufficient for PD-L1 testing by the central lab.

Participants will be assessed for complete study eligibility prior to randomization as specified in Section 2.

The Screening stage ends with either confirmation of full eligibility and randomization for the participant or with the confirmation that the participant is a screen failure. This study permits the re-enrollment of a participant that has discontinued the study as a pre-treatment failure prior to randomization. If re-enrolled, the participant must be re-consented. A new participant identification number will be assigned by IRT at the time of re-enrollment.

Treatment stage:

CA2099ER Global Revised Protocol 01

Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B. Participants previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinical planned events, per protocol.

The Treatment stage begins when the randomization call is made into the IRT. The participant is randomly assigned to 1 of the 2 treatment arms as noted in the study schematic above.

- Arm A (Doublet): Nivolumab 240 mg flat dose IV Q2W + Cabozantinib 40 mg PO QD
    - Nivolumab to be continued until disease progression or unacceptable toxicity with maximum treatment of 2 years from the first dose in Cycle 1
    - Cabozantinib to be continued until disease progression or unacceptable toxicity
- Arm C: Sunitinib 50 mg PO QD for 4 weeks, followed by 2 weeks off, per cycle. Cycles to be continued until progression or unacceptable toxicity

- Note - Randomization to Arm B stops with implementation of approved CA2099ER Global Revised Protocol 01. Treatment B (below) continues only for participants randomized to Arm B prior to implementation of Global Revised Protocol 01
    - Arm B (Triplet): Nivolumab 3mg/kg IV + Ipilimumab 1 mg/kg IV, both Q3W x 4 doses + Cabozantinib 40 mg PO QD
        - Then Nivolumab 240 mg flat dose IV Q2W + Cabozantinib 40 mg PO QD
        - Nivolumab to be continued until disease progression, unacceptable toxicity, or a maximum of 2 years from the first dose in Cycle 1
        - Cabozantinib to be continued until progression or unacceptable toxicity

Study treatment must begin within 3 days (72 hours) of randomization. Participants in Arm A will continue nivolumab until progression, unacceptable toxicity, withdrawal of consent, or a maximum of 2 years from the first dose in Cycle 1, whichever occurs first. Cabozantinib (Arm A) may be continued until progression, unacceptable toxicity, or withdrawal of consent, whichever occurs first, and may extend beyond 2 years from the first dose in Cycle 1. See Table 7.1-1 and

Table 7.1-2 for the dosing schedule. Study drugs may be delayed for toxicity (See Section 7.4.1). Treatment may be continued beyond investigator-assessed progression if the investigator confirms that the participant meets the criteria specified in Section 8.1.4.

A negative pregnancy test should be documented within 24 hours prior to the initial dose of the investigational product and then performed every 4 weeks ± 7 days during treatment. On-study laboratory assessments should be drawn within 72 hours prior to dosing and will be assessed at the local laboratory.

Tumor assessments will occur in accordance with Section 2 and Section 9.1.2 until progression has been assessed by the investigator **and** confirmed by the blinded independent central review (BICR). Each site must submit scans on a rolling basis, preferably within 7 days of image acquisition, to a third-party vendor for BICR. If progression is assessed by the investigator, the site will inform the radiology vendor so that the BICR assessment of progression can be performed. The BICR assessment of progression will be completed, and the results provided to the site, within approximately 14 days, as specified in Section 9.1.2.

PK and immunogenicity samples will be collected according to Table 9.5-1 and Table 9.5-2. Adverse event assessments should be documented at each clinic visit.

Quality of Life will be assessed using the National Comprehensive Cancer Network (NCCN) Functional Assessment of Cancer Therapy - Kidney Symptom Index (FKSI-19) and EuroQoL's EQ-5D-3L. These questionnaires should be completed according to Section 2.

The Treatment stage ends when the participant is discontinued from study therapy.

Follow-up stage: The Follow-up stage begins when the decision to discontinue a participant from study therapy is made (no further treatment with study therapy). Participants must be followed for at least 100 days after last dose of study treatment. Follow-up visit #1 (FU1) should occur 30 days from the last dose (+/- 7) days or can be performed on the date of discontinuation if that date is greater than 42 days from last dose. Follow-up visit #2 (FU2) occurs approximately 100 days (+/- 7 days) from last dose of study drug. Both Follow Up visits should be conducted in person. AEs will be followed until the toxicities resolve, return to baseline, or are deemed irreversible. The FKSI-19 and EQ-5D-3L will be completed as described in Section 2.

Participants who discontinued study treatment without BICR confirmed radiographic progression will continue to have tumor assessments performed according to the frequency described in Sections 2 and 9.1.2, even if new anti-tumor therapy has been initiated. If progression is assessed by the investigator, the site will inform the radiology vendor, so that the BICR assessment of progression can be performed. Participants whose progression is not confirmed by the BICR will be required to continue tumor assessments (if clinically feasible) according to the protocol-specified schedule or sooner if clinically indicated until BICR confirms progression on a subsequent tumor assessment (See Section 9.1.2.2 for additional details).

After the Follow-up 2 Visit, all participants will be followed for overall survival status every 3 months (+/- 14 days) until death, withdrawal of consent, loss to follow-up, or end of study. Survival status can be ascertained in person or by telephone contact. If new anti-tumor therapy is

initiated for either progression or a secondary malignancy at any time during this period, this and all other pertinent data obtained should be recorded on the appropriate Case Report Form (CRF).

5.1.1 Data Monitoring Committee and Other External Committees

To provide independent oversight of safety, efficacy, and study conduct, a data monitoring committee (DMC) will be instituted. The DMC will meet regularly to ensure that participant safety is carefully monitored, including a safety assessment after the first 30 participants are randomized including participants in Arm B prior to global Revised Protocol 01 and are followed for at least 6 weeks and then again after the first 75 participants are randomized and followed for at least 6 weeks. The DMC will convene additional ad hoc meetings if necessary. Following each meeting, the DMC will recommend continuation, modification, or discontinuation of the study based on observed toxicities. The DMC will also review the interim analysis results and inform BMS whether stopping criteria for superiority are met at that time. A separate DMC charter will describe the activities of this committee in more detail.

Blinded independent central review (BICR) will be utilized in this study for determination of PFS endpoint. The BICR will review all available tumor assessment scans for all randomized participants. Details of BICR responsibilities and procedures will be specified in the BICR charter.

5.2 Number of Participants

Approximately 850 participants will be enrolled in order to randomize approximately 638 participants (319 per Arm A and per Arm C). The number of randomized participants with favorable risk disease will be capped at approximately 25% (160) participants.

5.3 End of Study Definition

The start of the trial is defined as the first visit for the first participant screened. End of trial is defined as the last visit or scheduled procedure shown in the Schedule of Activities for the last participant. Study completion is defined as the final date on which data for the endpoint of overall survival was or is expected to be collected, if this is not the same.

5.4 Scientific Rationale for Study Design

Please refer to Section 3.1 for details of the rationale for evaluating the combination of nivolumab with cabozantinib (doublet regimen) in this study.

5.4.1 Rationale for Open-Label Design

The trial will have an open-label design with a blinded independent central review. Different dosing schedules and use of both IV infusions and oral medications across the 2 different arms make blinding the trial impractical. The complexity of including multiple visits for placebo infusions are burdensome for this participant population. Nivolumab is associated with some toxicities that are also common to cabozantinib and sunitinib (eg, diarrhea, hepatotoxicity), and the management of these common toxicities are different for immuno-oncology agents (eg, typically requiring steroids) vs TKIs (eg, typically requiring dose delay and reductions). Therefore, an open-label trial is preferable for participant safety as it allows the optimal management of toxicities.

5.4.2 Rationale for Choice of Primary Endpoint

The primary endpoint to be evaluated is PFS per BICR in all randomized participants. PFS has been an acceptable endpoint to support regulatory approvals in first-line mRCC. Sorafenib, sunitinib, bevacizumab, and pazopanib were approved based on an increase in PFS. While OS remains the "gold standard" for oncology trials, as more active agents gain approval in RCC, subsequent treatment with these agents confounds the effect of first-line treatment on OS. Therefore, PFS remains an important indicator of clinical benefit in first-line mRCC and has been chosen as the primary endpoint for this trial.

The primary efficacy analysis population includes all randomized participants, including those with favorable, intermediate, or poor risk disease per IMDC prognostic criteria. Favorable risk will be capped at 25% to represent the normal frequency of favorable risk group in mRCC. Several TKIs were approved as the first line treatment for mRCC in all risk groups. The combination regimen of a TKI is likely to benefit across all IMDC risk groups. Nivolumab and cabozantinib should be evaluated in this population for the primary analysis versus Arm C (sunitinib) as the comparator.

5.5 Justification for Dose

5.5.1 Dosing for Sunitinib Monotherapy in Arm C

Dosing for sunitinib monotherapy in Arm C uses the standard dosing schedule found in the sunitinib prescribing information.[33]

5.5.2 Dosing for Cabozantinib Doublet and Triplet Therapy in Arm A and Arm B

Dosing for nivolumab combined with cabozantinib (doublet regimen) was determined based on a phase 1b/2 trial (NCT02496208) conducted with the combination at NCI/NIH (see Section 3.2.1.4).

Note: Arm B (triplet regimen) in the original protocol has been removed from this Revised Protocol 01. It is more appropriate to study nivolumab ipilimumab containing regimen (e.g. Arm B, triplet regimen) versus a control arm of nivolumab and ipilimumab.

5.5.3 Dosing Rationale for Nivolumab 240 mg flat dose Q2W

A flat dose of nivolumab 240 mg every 2 weeks will be administered in Arm A and in the maintenance phase of Arm B in combination with cabozantinib. Nivolumab 3mg/kg once every 2 weeks has been shown to be similar to 240 mg flat dose every 2 weeks, and is currently FDA approved, based on clinical data and modeling and simulation approaches using population PK (PPK) and exposure-response analyses of data from studies in multiple tumor types (melanoma, non-small-cell lung cancer [NSCLC], renal cell carcinoma [RCC]) and urothelial carcinoma where body weight normalized dosing (mg/kg) has been used.

PPK analyses have shown that the PK of nivolumab is linear with proportional exposure over a dose range of 0.1 to 10 mg/kg, and no differences in PK across ethnicities and tumor types were observed. Nivolumab clearance and volume of distribution were found to increase as the body

weight increases, but less than proportionally with increasing weight, indicating that mg/kg dosing represents an over-adjustment for the effect of body weight on nivolumab PK. The PPK model developed using data from NSCLC participants was updated, using data from participants investigating nivolumab in the treatment of melanoma, NSCLC, and RCC. In this dataset, the median (minimum to maximum) weight was 77 kg (35 to 160 kg) and thus, an approximately equivalent dose of 3 mg/kg for an 80 kg participant, nivolumab 240 mg Q2W was selected for future studies. To predict relevant summary exposures of nivolumab 240 mg Q2W, the PPK model was used to simulate nivolumab 3 mg/kg Q2W and 240 mg Q2W. In the simulations, the simulated participant populations consisted of 1000 participants per treatment arm randomly sampled from aforementioned pooled database of cancer participants. Because no differences in PK were noted across ethnicities and tumor types, these simulated melanoma, NSCLC and RCC data will be applicable to participants with other tumor types. The simulated measure of exposure of interest, time-averaged concentrations (Cavgss) for 240 mg Q2W were predicted to be similar for all participants in reference to 80 kg participants receiving 3 mg/kg Q2W.

Additionally, nivolumab is safe and well tolerated up to 10 mg/kg Q2W dose level. Adverse events have been consistent across tumor types following monotherapy and have not demonstrated clear dose-response or exposure-response relationships. Additionally, the simulated median and 95th prediction interval of nivolumab summary exposures across a wide body weight range (35 - 160 kg) are predicted to be maintained below the corresponding observed highest exposure experienced in nivolumab, ie, 95th percentile following nivolumab 10 mg/kg Q2W from clinical study CA209003. Thus, while participants in the lower body weight ranges would have greater exposures than 80 kg participants, the exposures are predicted to be within the range of observed exposures at doses (up to 10 mg/kg Q2W) used in the nivolumab clinical program, and are not considered to put participants at increased risk. Additionally, for participants with greater body weights, the simulated ranges of exposures are also not expected to affect efficacy, because the exposures predicted following administration of a 240 mg Q2W are on the flat part of the exposure-response curves for previously investigated tumors, melanoma and NSCLC. Given the similarity of nivolumab PK across many tumor types, including RCC, and the similar exposures predicted following administration of 240 mg flat dose compared to 3 mg/kg, it was shown that the safety and efficacy profile of 240 mg nivolumab will be similar to that of 3 mg/kg nivolumab.

5.5.4 Rationale for Dosing Duration for Nivolumab

The optimal duration of immunotherapy is an important question and continues to be investigated. Clinical trials across different tumors types in the nivolumab and ipilimumab development program indicate that most of the responses occur early, with a median time to response of 2-4 months, and emerging data suggests that benefit can be maintained in the absence of continued treatment. A recent analysis in a melanoma study suggests the majority of patients who discontinue nivolumab and/or ipilimumab for toxicity maintain disease control in the absence of further treatment.[44] Furthermore, a limited duration of ipilimumab, including only 4 induction doses, resulted in long term survival in patients with metastatic melanoma, with a sustained plateau in survival starting around 2 years after the start of treatment.[45]

Accumulating data suggest that 2 years of PD-1 checkpoint inhibitor treatment may be sufficient for long term benefit. CA209003, a dose-escalation cohort expansion trial evaluating the safety and clinical activity of nivolumab in patients with previously treated advanced solid tumors (including 129 subjects with NSCLC), specified a maximum treatment duration of 2 years. Among 16 subjects with non-small cell lung cancer (NSCLC) who discontinued nivolumab after completing 2 years of treatment, 12 subjects were alive >5 years and remained progression-free without any subsequent therapy (2) In the CA209003 NSCLC cohort, the overall survival (OS) curve begins to plateau after 2 years, with an OS rate of 25% at 2 years and 18% at 3 years.[46] These survival outcomes are similar to Phase 3 studies in previously treated NSCLC, in which nivolumab treatment was continued until progression or unacceptable toxicity (2 year OS rates of 23% and 29%, and 3-year OS rates of 16%-18% for squamous and non-squamous NSCLC respectively).[47]

Similar results have been reported in clinical studies of pembrolizumab, another PD-1 inhibitor. Keynote-010 was a randomized Phase 3 trial of pembrolizumab (at either 2 mg/kg or 10 mg/kg every 3 weeks) versus docetaxel in subjects with previously treated, PD-L1-positive, advanced NSCLC which specified a maximum treatment duration of 2 years for pembrolizumab. OS was significantly longer with both pembrolizumab 2 mg/kg (HR 0.72, $p = 0.00017$) and pembrolizumab 10 mg/kg (HR 0.60, $p < 0.00001$) compared to docetaxel, with an OS plateau developing beyond 2 years in both pembrolizumab arms. Among 690 patients who received pembrolizumab, 47 patients completed 2 years of pembrolizumab and stopped treatment. Most were able to maintain their response, including those with stable disease, with only 2 patients (4%) having confirmed progression after stopping at 2 years.[48]

Keynote-006 was a randomized phase 3 study of pembrolizumab versus ipilimumab in patients with advanced melanoma, which also specified a maximum 2 year duration of pembrolizumab treatment. 104 (19%) of 556 patients randomized to pembrolizumab completed 2 years of treatment. With a median follow-up of 9 months after completion of pembrolizumab, the estimated risk of progression or death was 9% in these patients.[49]

Taken together, these data suggest that treatment beyond 2 years is unlikely to confer additional clinically meaningful benefit and that the risk of progression after discontinuing treatment at 2 years is low.

In contrast, a shorter duration of nivolumab of only 1 year was associated with increased risk of progression in previously treated patients with NSCLC, suggesting that treatment beyond 1 year is likely needed. In CA209153, patients with previously treated advanced NSCLC who completed 1 year of nivolumab therapy were randomized to either continue or stop treatment, with the option of retreatment upon progression. Among 163 patients still on treatment at 1 year and without progression, those who were randomized to continue nivolumab had significant improvement in progression-free survival (PFS) compared to those who were randomized to stop treatment, with median PFS (post-randomization) not reached vs 10.3 months, respectively; HR=0.42 (95% CI, 0.25 to 0.71). With a median follow-up of 14.9 months post-randomization, there also was a trend for patients on continued treatment to live longer (OS HR = 0.63 [95% CI: 0.33, 1.20]). Of note,

the PFS curves in both groups plateau approximately 1 year after randomization (i.e., 2 years after treatment initiation), suggesting that there may be minimal benefit in extending treatment beyond a total of 2 years.[50]

Collectively, these data suggest that there is minimal if any benefit derived from continuing I-O treatment beyond two years in advanced tumors. However, even though immunotherapy is well tolerated, patients will be at risk for additional toxicity with longer term treatment. Therefore, in this study, treatment will be given for a maximum of 2 years from the start of study treatment.

6 STUDY POPULATION

For entry into the study, the following criteria MUST be met.

6.1 Inclusion Criteria

1) Signed Written Informed Consent

a) Participants must have signed and dated an IRB/IEC approved written informed consent form in accordance with regulatory and institutional guidelines. This must be obtained before the performance of any protocol related procedures that are not part of normal participant care.

b) Participants must be willing and able to comply with scheduled visits, treatment schedule, and laboratory testing.

2) Type of Participant and Target Disease Characteristics

a) Histological confirmation of RCC with a clear-cell component, including participants who may also have sarcomatoid features

b) Advanced (not amenable to curative surgery or radiation therapy) or metastatic (AJCC Stage IV) RCC

c) No prior systemic therapy for RCC with the following exception:

i) One prior adjuvant or neoadjuvant therapy for completely resectable RCC if such therapy did not include an agent that targets VEGF or VEGF receptors and if recurrence occurred at least 6 months after the last dose of adjuvant or neoadjuvant therapy.

d) Karnofsky Performance Status (KPS) ≥ 70% (See Appendix 7)

e) Measurable disease as per RECIST v1.1 per investigator. (See Appendix 8)

f) Either a formalin-fixed, paraffin-embedded (FFPE) tissue block or unstained tumor tissue sections, preferably obtained within 3 months but no more than 12 months prior to enrollment, with an associated pathology report, must be submitted to the central laboratory during screening. Biopsy should be excisional, incisional or core needle. Fine needle aspiration is unacceptable for submission. In participants who received prior adjuvant or neoadjuvant therapy, the tumor sample must have been obtained after completion of adjuvant or neoadjuvant therapy. Upon receipt of the tumor sample, the central lab will determine PD-L1 expression level by IHC testing (see Section 9.8.2). In order to be randomized in the IRT system, a participant must be classified as PD-L1 expression ≥ 1%, PD-L1 expression < 1%, or PD-L1 expression indeterminate. Sites will be informed when the submitted tumor sample is insufficient for PD-L1 testing by the central lab.

g) Participants with favorable, intermediate and poor risk categories will be eligible for the study. To be eligible for the Intermediate and Poor-Risk cohort, at least one of the following prognostic factors as per International Metastatic RCC Database Consortium (IMDC) must be present: (See Appendix 6)

i) KPS equal to 70%

ii) Less than 1 year from initial diagnosis (including original localized disease if applicable) to randomization

iii) Hemoglobin < lower limit of normal (LLN)

iv) Corrected calcium concentration > 10 mg/dL

v) Absolute neutrophil count > ULN

vi) Platelet count > ULN

If none of the above factors are present, participants are only eligible for the favorable-risk cohort. Approximately 160 favorable risk participants will be randomized. Enrollment of favorable risk participants may be closed earlier than enrollment for intermediate and poor risk participants.

3) Age and Reproductive Status

a) Males and Females, ages 18 or age of majority

b) Women of childbearing potential (WOCBP) must have a negative serum or urine pregnancy test (minimum sensitivity 25 IU/L or equivalent units of HCG) within 24 hours prior to the start of study treatment

c) Women must not be breastfeeding

d) Women of childbearing potential (WOCBP) must agree to follow instructions for method(s) of contraception for the duration of study treatment and 5 months after the last dose of study treatment (ie, 30 days [duration of ovulatory cycle] plus the time required for the investigational drug to undergo approximately five half-lives).

e) Males who are sexually active with WOCBP must agree to follow instructions for method(s) of contraception for the duration of study treatment and 7 months after the last dose of study treatment (ie, 90 days [duration of sperm turnover] plus the time required for the investigational drug to undergo approximately five half-lives).

f) Azoospermic males are exempt from contraceptive requirements. WOCBP who are continuously not heterosexually active are also exempt from contraceptive requirements, and still must undergo pregnancy testing as described in this section.

Investigators shall counsel WOCBP, and male participants who are sexually active with WOCBP, on the importance of pregnancy prevention and the implications of an unexpected pregnancy. Investigators shall advise on the use of highly effective methods of contraception (Appendix 4) which have a failure rate of < 1% when used consistently and correctly.

6.2 Exclusion Criteria

1) Medical Conditions

a) Any active CNS metastases. Participants with treated, stable CNS metastases for at least 1 month are eligible as long as they meet the following criteria:

Treated CNS metastases are defined as having no ongoing requirement for corticosteroids for at least 2 weeks prior to randomization and no evidence of progression or hemorrhage after treatment completed at least 1 month prior to randomization, as ascertained by clinical examination and brain imaging (MRI or CT). (Stable dose of anticonvulsants is allowed). Treatment for CNS metastases may include whole brain radiotherapy, radiosurgery (eg, RS, gamma knife, LINAC, or equivalent) or a combination as deemed appropriate by the treating physician. Participants with CNS metastases treated by neurosurgical resection or brain biopsy performed within 1 month prior to randomization are not eligible. Baseline imaging of the brain is required within 28 days prior to randomization.

b) Any active, known or suspected autoimmune disease. Participants with type I diabetes mellitus, hypothyroidism only requiring hormone replacement, skin disorders (such as vitiligo, psoriasis, or alopecia) not requiring systemic treatment, or conditions not expected to recur in the absence of an external trigger (e.g. celiac disease) are permitted to enroll.

c) Any condition requiring systemic treatment with either corticosteroids (> 10 mg daily prednisone equivalent) or other immunosuppressive medications within 14 days of randomization. Inhaled or topical steroids, and adrenal replacement steroid doses > 10 mg daily prednisone equivalent, are permitted in the absence of active autoimmune disease.

d) Prior malignancy active within the previous 3 years except for locally curable cancers that have been apparently cured, such as basal or squamous cell skin cancer, superficial bladder cancer, or carcinoma in situ of the prostate, cervix, or breast

e) Any tumor invading the SVC or other major blood vessels

f) Any tumor invading the GI tract or any evidence of endotracheal or endobronchial tumor within 30 days prior to randomization

g) Known history of positive test for human immunodeficiency virus (HIV) or known acquired immunodeficiency syndrome (AIDS). NOTE: Testing for HIV must be performed at sites where mandated locally (sites in Germany, see Appendix 12).

h) Known medical condition (eg, a condition associated with diarrhea or acute diverticulitis, aortic aneurysm, aortic dissection) that, in the investigator's opinion, would increase the risk associated with study participation or study drug administration or interfere with the interpretation of safety results

i) History of abdominal fistula, gastrointestinal perforation, intra-abdominal abscess, bowel obstruction, or gastric outlet obstruction within the past 6 months prior to randomization

j) Impairment of gastrointestinal function or gastrointestinal disease that may significantly alter the absorption of cabozantinib or sunitinib (eg, malabsorptive disorder, ulcerative disease, uncontrolled nausea, vomiting, diarrhea, or small bowel resection)

k) Serious, non-healing wound or ulcer within 30 days prior to randomization

l) Evidence of active bleeding or bleeding susceptibility; or medically significant hemorrhage within prior 3 months prior to randomization

m) Uncontrolled adrenal insufficiency

n) History of cerebrovascular accident (CVA) including transient ischemic attack within the past 6 months prior to randomization

o) History of deep vein thrombosis (DVT) or pulmonary embolism (PE) within past 6 months prior to randomization unless stable, asymptomatic, and treated with low molecular weight heparin (LMWH) for at least 3 weeks prior to randomization

p) Any unstable cardiac arrhythmia within 6 months prior to randomization

q) Prolongation of the Fridericia corrected QT (QTcF) interval defined as > 450 msec for males and > 470 msec for females, where QTcF = QT / $3\sqrt{RR}$ with triplicate measurements

r) Poorly controlled hypertension (defined as systolic blood pressure (SBP) of > 150 mmHg or diastolic blood pressure (DBP) of > 90 mmHg), despite antihypertensive therapy

s) History of any of the following cardiovascular conditions within 6 months of randomization: cardiac angioplasty or stenting, myocardial infarction, unstable angina, coronary artery by-pass graft surgery, symptomatic peripheral vascular disease, class III or IV congestive heart failure (CHF), as defined by the New York Heart Association (NYHA)

t) Any radiologic or clinical evidence of pancreatitis within 30 days prior to randomization

u) Inability to swallow oral medications

2) Prior/Concomitant Therapy

a) Prior treatment with VEGF, MET, AXL, KIT, or RET targeted therapy (including, but not limited to, sunitinib, pazopanib, axitinib, tivozanib, sorafenib, lenvatinib, bevacizumab, and cabozantinib).

b) Prior treatment with an anti-PD-1, anti-PD-L1, anti-PD-L2, anti-CD137, or anti-CTLA-4 antibody, or any other antibody or drug specifically targeting T-cell co-stimulation or checkpoint pathways.

c) Concomitant strong CYP3A4 inducers or inhibitors within 14 days prior to randomization. Please refer to Appendix 9 for a list of common strong CYP3A4 inducers and inhibitors.

d) Concomitant treatment, in therapeutic doses, with anticoagulants such as warfarin or warfarin-related agents, thrombin or Factor Xa inhibitors. Aspirin (up to 325 mg/day) and prophylactic and therapeutic low molecular weight heparin (LMWH) are permitted.

e) Major surgery less than 6 weeks, nephrectomy less than 4 weeks, prior to randomization, with complete wound healing and no ongoing post-operative complications.

f) Any of the following prior radiotherapy procedures:

i) Radiotherapy to the thoracic cavity or abdomen within 4 weeks prior to randomization

ii) Radiotherapy to bone lesions within 2 weeks prior to randomization

iii) Radiotherapy to any other site within 4 weeks prior to randomization

NOTE: In all cases, there must be complete recovery and no ongoing complications from prior radiotherapy.

g) Treatment with botanical preparations (eg herbal supplements or traditional Chinese medicines) intended for general health support or to treat the disease under study within 2 weeks prior to randomization.

h) Participants who have received a live/attenuated vaccine within 30 days of first treatment.

3) Physical and Laboratory Test Findings

a) Ejection fraction ≤ 50% on screening echocardiogram or MUGA

b) WBC < 2000/μL

c) Neutrophils < 1500/μL
d) Platelets < 100 x 10^3/μL
e) Hemoglobin < 9.0 g/dL (support with transfusion is acceptable)
f) Serum creatinine > 1.5 x ULN unless calculated creatinine clearance ≥ 40 mL/min (using the Cockcroft-Gault formula)
g) AST/ALT > 3.0 x ULN
h) Total bilirubin > 1.5 x ULN (except participants with Gilbert Syndrome who must have a total bilirubin level of < 3.0 x ULN)
i) Urine protein/creatinine ratio (UPCR) > 1.5, unless 24-hour urine protein is ≤ 1.5 g
j) INR > 1.5
k) Any positive test result for hepatitis B virus or hepatitis C virus indicating presence of virus, eg, hepatitis B surface antigen (HBsAg) positive, or hepatitis C antibody (anti-HCV) positive (except if HCV-RNA negative)

4) Allergies and Adverse Drug Reaction

a) History of allergy or hypersensitivity to study drug components
b) No history of severe hypersensitivity to a monoclonal antibody

5) Other Exclusion Criteria

a) Prisoners or participants who are involuntarily incarcerated. (Note: under specific circumstances, and only in countries where local regulations permit, a person who has been imprisoned may be included as a participant. Strict conditions apply and Bristol-Myers Squibb approval is required.)
b) Participants who are compulsorily detained for treatment of either a psychiatric or physical (eg, infectious disease) illness

Eligibility criteria for this study have been carefully considered to ensure the safety of the study participants and that the results of the study can be used. It is imperative that participants fully meet all eligibility criteria.

6.3 Lifestyle Restrictions

6.3.1 Meals and Dietary Restrictions

For participants in Arms A (and Arm B participants enrolled prior to Revised Protocol 01), cabozantinib should not be taken with grapefruit/grapefruit juice or Seville oranges.

For participants in Arm C, sunitinib should not be taken with grapefruit/grapefruit juice.

6.4 Screen Failures

Screen failures are defined as participants who consent to participate in the clinical study but are not subsequently randomized. A minimal set of screen failure information is required to ensure transparent reporting of screen failure participants, to meet the Consolidated Standards of Reporting Trials (CONSORT) publishing requirements, and to respond to queries from regulatory authorities. Minimal information includes date of consent, demography, screen failure details, eligibility criteria, and any serious AEs.

6.4.1 Retesting During Screening or Lead-In Period

Participant Re-enrollment: This study permits the re-enrollment of a participant that has discontinued the study as a pre-treatment failure (ie, participant has not been randomized / has not been treated). If re-enrolled, the participant must be re-consented.

Retesting of laboratory parameters and/or other assessments within any single Screening or Lead-in period will be permitted (in addition to any parameters that require a confirmatory value).

The most current result prior to randomization is the value by which study inclusion will be assessed, as it represents the participant's most current, clinical state.

Laboratory parameters and/or assessments that are included in Table 2-1 Screening Procedural Outline may be repeated in an effort to find all possible well-qualified participants. Consultation with the BMS Medical Monitor may be needed to identify whether repeat testing of any particular parameter is clinically relevant.

7 TREATMENT

Study treatment is defined as any investigational treatment(s), marketed product(s), placebo or medical device intended to be administered to a study participant according to the study randomization or treatment allocation.

Study treatment includes both Investigational [Medicinal] Product (IP/IMP) and Non-investigational [Medicinal] Product (Non-IP/Non-IMP) and can consist of the following:

An investigational product, also known as investigational medicinal product in some regions, is defined a pharmaceutical form of an active substance or placebo being tested or used as a reference in a clinical study, including products already with a marketing authorization but used or assembled (formulated or packaged) differently than the authorized form, or used for an unauthorized indication, or when used to gain further information about the authorized form.

Other medications used as support or escape medication for preventative, diagnostic, or therapeutic reasons, as components of the standard of care for a given diagnosis, may be considered as non-investigational products.

Table 7-1: Study treatments for CA2099ER

Product Description / Class and Dosage Form	Potency	IP/ Non-IMP	Blinded or Open Label	Packaging / Appearance	Storage Conditions (per label)
BMS-936558-01 (Nivolumab) Solution for Injection*	100 mg (10 mg/mL)	IP	Open Label	Vial (multiple vials per carton)	Store at 2° - 8 °C. Protect from light and freezing
Ipilimumab Solution for Injection (see Note)	200 mg (5mg/mL)	IP	Open Label	Vial (multiple vials per carton)	Store at 2° - 8 °C. Protect from light and freezing
Cabozantinib Tablet	20 mg	IP	Open Label	Tablets in a bottle	Refer to storage conditions on container label
Sunitinib Malate Capsule**	12.5 mg	IP	Open Label	Capsules in various packaging configurations	Refer to storage conditions on container label/package insert

*May be labeled as "BMS-936558-01" or "Nivolumab"

** Sunitinib may be obtained by the investigational sites in certain countries as local commercial product (which may be available as a different potency/package size than listed above) if local regulations allow and agreed to by BMS.

Note: Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B. Participants previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinically planned events, per protocol

Premedications or medications used to treat infusion-related reactions should be sourced by the investigative sites if available and permitted by local regulations. Solutions used as diluent (ie, 0.9% Sodium Chloride Injection or 5% Dextrose Injection) should also be sourced by investigative sites if available and permitted by local regulations.

7.1 Treatments Administered

The selection and timing of dose for each participant is presented in Table 7.1-1 and the dosing scheduled by cycle is presented in Table 7.1-2.

Table 7.1-1: Selection and Timing of Dose

Arm	Study Treatment	Dosage level(s) and Formulation	Frequency of Administration	Route of Administration
A (Doublet)	Nivolumab	240 mg IV	Every 2 weeks	IV
	Cabozantinib	40 mg (20 mg tablets)	Once daily (QD)	PO
B (Triplet) See Note	Nivolumab	3 mg/kg IV for 4 doses then 240 mg IV	Every 3 weeks (Q3W) for 4 doses then every 2 weeks (Q2W)	IV
	Ipilimumab	1 mg/kg IV for 4 doses	Every 3 weeks (Q3W) for 4 doses	IV
	Cabozantinib	40 mg (20 mg tablets)	Once daily (QD)	PO
C	Sunitinib	50 mg (12.5 mg capsules)	A 6 week cycle, consisting of once daily (QD) regimen for 4 weeks followed by no treatment for 2 weeks.	PO

Abbreviations: IV=intravenous; PO= by mouth.

Note: Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B. Participants previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinically planned events, per protocol

Table 7.1-2: Dosing Schedule for CA2099ER

Arm	Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5 and subsequent cycles
Arm A Nivolumab + Cabozantinib (each cycle = 2 weeks)	Nivolumab 240 mg IV on Day 1 + Cabozantinib 40 mg PO QD	Nivolumab 240 mg IV on Day 1 + Cabozantinib 40 mg PO QD	Nivolumab 240 mg IV on Day 1 + Cabozantinib 40 mg PO QD	Nivolumab 240 mg IV on Day 1 + Cabozantinib 40 mg PO QD	Nivolumab 240 mg IV on Day 1 + Cabozantinib 40 mg PO QD
Arm B Nivolumab + Ipilimumab + Cabozantinib (see Note) (Cycle 1 to 4 = 3 weeks, Cycle 5 and subsequent cyces = 2 weeks)	Nivolumab 3 mg/kg IV on Day 1 + Ipilimumab 1 mg/kg IV on Day 1 + Cabozantinib 40 mg PO QD	Nivolumab 3 mg/kg IV on Day 1 + Ipilimumab 1 mg/kg IV on Day 1 + Cabozantinib 40 mg PO QD	Nivolumab 3 mg/kg IV on Day 1 + Ipilimumab 1 mg/kg IV on Day 1 + Cabozantinib 40 mg PO QD	Nivolumab 3 mg/kg IV on Day 1 + Ipilimumab 1 mg/kg IV on Day 1 + Cabozantinib 40 mg PO QD	Nivolumab 240 mg IV on Day 1 + Cabozantinib 40 mg PO QD

Table 7.1-2: Dosing Schedule for CA2099ER

Arm	Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5 and subsequent cycles
Arm C Sunitinib (each cycle = 6 weeks, treatment for 4 weeks then 2 weeks off each cycle)	Sunitinib 50 mg PO QD x 4 weeks	Sunitinib 50 mg PO QD x 4 weeks	Sunitinib 50 mg PO QD x 4 weeks	Sunitinib 50 mg PO QD x 4 weeks	Sunitinib 50 mg PO QD x 4 weeks

Abbreviations: IV=intravenous; PO= by mouth; QD= once daily

Note: Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B. Participants previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinically planned events, per protocol

Participants should begin study treatment within 3 days (72 hours) of randomization.

For **Arm A**, participants should receive nivolumab at a dose of 240 mg as an approximately 30 minute infusion on Day 1 of each treatment cycle until progression, unacceptable toxicity, withdrawal of consent, completion of 24 months of treatment (from the first dose on Cycle 1), or the end of the study, whichever occurs first. The first cabozantinib dose should be given in the evening on Day 1, Cycle 1 (after the Cycle 1 nivolumab dose).

For **Arm B (Note: only applicable to participants who randomized to Arm B prior to** Revised Protocol 01**)**, when nivolumab and ipilimumab are to be administered on the same day, nivolumab is to be administered first. Nivolumab infusion (approximately 30 minutes) must be promptly followed by a flush of diluent to clear the line of nivolumab before starting the ipilimumab infusion. The second infusion will always be the ipilimumab study drug (approximately 30 minutes infusion) and will start after the infusion line has been flushed, filters changed and participant has been observed to ensure no infusion reaction has occurred. The time in between infusions is expected to be at least 30 minutes (from the end of the nivolumab infusion to the start of the ipilimumab infusion). The first cabozantinib dose should be given in the evening on Day 1, Cycle 1 (after the Cycle 1 nivolumab and ipilimumab doses).

Starting with Cycle 5 in Arm B, participants should receive nivolumab at a dose of 240 mg as a 30 minute infusion on Day 1 of each treatment cycle until progression, unacceptable toxicity, withdrawal of consent, completion of 24 months of treatment (from the first dose on Cycle 1), or the end of the study, whichever occurs first.

Dosing calculations for nivolumab and ipilimumab should be based on the body weight assessed at baseline. It is not necessary to re-calculate subsequent doses if the participant weight is within 10% of the weight used to calculate the previous dose. All doses should be rounded up or to the nearest milligram per institutional standard.

There will be no dose escalations or reductions of nivolumab or ipilimumab allowed. Participants may be dosed with nivolumab no less than 12 days from the previous dose during q2w cycles.

Participants may be dosed with nivolumab and ipilimumab no less than 19 days from the previous dose during q3w cycles (ie, Cycles 1-4 in Arm B). Participants in Arm B may be dosed with the first nivolumab maintenance dose (Cycle 5) no less than 19 days from the previous nivolumab and ipilimumab doses. Premedications are not recommended for the first dose of nivolumab or ipilimumab.

Participants should be carefully monitored for infusion reactions during nivolumab and/or ipilimumab administration. If an acute infusion reaction is noted, participants should be managed according to Section 7.4.4.

Doses of nivolumab and/or ipilimumab may be interrupted, delayed, or discontinued depending on how well the participant tolerates the treatment. Dosing visits are not skipped, only delayed. The only exception is Cycle 4 in Arm B, which may be skipped (omitted) only for the reasons specified in Section 7.4.3.1and Section 8.1.1 below.

7.1.1 Nivolumab

Nivolumab Injection, 100 mg/10 mL (10 mg/mL) is to be administered as an IV infusion through a 0.2-micron to 1.2-micron pore size, low-protein binding in-line filter at the protocol-specified doses. It is not to be administered as an IV push or bolus injection. Please refer to the Investigational Brochure/pharmacy manual for further details regarding storage, preparation and administration. Care must be taken to assure sterility of the prepared solution as the product does not contain any antimicrobial preservative or bacteriostatic agent.

7.1.2 Ipilimumab (Participants Randomized to Arm B Prior to Revised Protocol 01)

For details regarding ipilimumab storage, preparation, and administration, please refer to the instructions in the ipilimumab IB and/or pharmacy manual.

Separate infusion bags and filters should be used when administering nivolumab and ipilimumab on the same day.

Care must be taken to assure sterility of the prepared solutions, since the drug product does not contain any antimicrobial preservatives or bacteriostatic agents.

7.1.3 Cabozantinib

Cabozantinib is taken orally on an empty stomach, preferably at bed time. Participants should fast (with the exception of water) for at least 2 hours before until 1 hour after each dose of cabozantinib. Cabozantinib tablets should not be crushed or chewed. Cabozantinib should not be taken with grapefruit/grapefruit juice or Seville oranges. Missed doses of cabozantinib should not be taken within 12 hours of the next dose.

Dispense cabozantinib tablets in their original containers.

7.1.4 Sunitinib

Sunitinib is taken orally without regard to meals. Participants are to avoid grapefruit/grapefruit juice while on treatment with sunitinib.

7.2 Method of Treatment Assignment

All participants will be centrally randomized using an Interactive Response Technology (IRT). Before the study is initiated, each user will receive log in information and directions on how to access the IRT. Specific instructions for using IRT will be provided to the investigational site in a separate document.

The investigator or designee will register the participant for enrollment by following the enrollment procedures established by BMS.

Once enrolled in IRT, enrolled participants that have met all eligibility criteria, including determination of PD-L1 expression in the tumor sample submitted to and evaluated by the central laboratory, will be ready to be randomized through the IRT prior to the start of study treatment administration for each participant. The site will record the treatment assignment on the applicable case report form, if required.

Study treatment will be dispensed at the study visits as listed in Schedule of Activities (Section 2).

7.3 Blinding

Not applicable as this is an open-label study; however, the specific treatment to be taken by a participant will be assigned using an Interactive Response Technology (IRT).

Treatment assignments will be released to the bioanalytical laboratory in order to minimize unnecessary analysis of samples.

BICR will remain blinded to treatment assignment.

7.4 Dosage Modification

CA2099ER Global Revised Protocol 01

Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B. Participants previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinical planned events, per protocol.

In Arm A and Arm B, the assessment of causality for any AE should be performed independently for each study drug in the combination regimen.

In the doublet treatment arm (Arm A), cabozantinib and nivolumab will be administered at Cycle 1. After Cycle 1, modifications in cabozantinib dosing (delay, reduction/escalation, and discontinuation) may occur as outlined in Section 7.4.1, Section 7.4.2, and Section 8. After Cycle 1, nivolumab may be modified as outlined in Section 7.4.1, Section 7.4.2, and Section 8.

In the triplet treatment arm (Arm B), cabozantinib, nivolumab, and ipilimumab will be administered at Cycle 1. After Cycle 1, modifications in cabozantinib dosing (delay, reduction/escalation, and discontinuation) may occur as outlined in Section 7.4.1, Section 7.4.2, and Section 8. After Cycle 1, nivolumab or ipilimumab may be modified as outlined in Section 7.4.1, Section 7.4.2, and Section 8.

In Arm C, sunitinib will be administered at Cycle 1. After Cycle 1, modifications in sunitinib dosing (delay, reduction/escalation, and discontinuation) may occur as outlined in Section 7.4.1, Section 7.4.2, and Section 8.

7.4.1 Dose Delay Criteria

Dose delay criteria for management of adverse events during nivolumab, ipilimumab, cabozantinib, or sunitinib treatment are outlined in this section.

Dosing of nivolumab (in Arm A) or nivolumab and ipilimumab (in Arm B) may be delayed without delay of cabozantinib dosing if toxicity is felt to be related to only nivolumab (Arm A) or nivolumab and ipilimumab (Arm B) and not related to cabozantinib. Conversely, dosing of cabozantinib may be delayed without delay of nivolumab dosing (in Arm A) or nivolumab and ipilimumab dosing (in Arm B) if toxicity is felt to be related to only cabozantinib and not related to nivolumab (Arm A) or nivolumab and ipilimumab (Arm B). However, if toxicity is considered related to all study drugs or if the investigator is unable to determine which study drug is the cause of the AE, then all study drugs in the combination should be delayed.

Participants who require dose delay should be re-evaluated weekly or more frequently if clinically indicated and resume dosing when re-treatment criteria are met (see Section 7.4.3).

7.4.1.1 Dose Delay Criteria for Nivolumab and Ipilimumab

In Arm A, nivolumab administration should be delayed, and in Arm B, both nivolumab and ipilimumab administration should be delayed, for any of the following:

- Any Grade ≥ 2 non-skin, drug-related AE, with the exception of fatigue
- Grade 2 drug-related creatinine, AST, ALT, and/or total bilirubin abnormalities
- Any Grade 3 skin, drug-related AE
- Any Grade 3 drug-related laboratory abnormality, with the following exceptions:
  - Grade 3 lymphopenia does not require dose delay
  - Grade 3 amylase or lipase abnormalities that are not associated with symptoms or clinical manifestations of pancreatitis do not require a dose delay.
  - Grade ≥ 3 AST, ALT or total bilirubin will require dose discontinuation (see section 8.1.1)
- Any adverse event, laboratory abnormality, or intercurrent illness which, in the judgment of the investigator, warrants delaying the dose of study medication.

During Cycles 1-4 in Arm B, both nivolumab and ipilimumab must be delayed at the same time.

Immuno-oncology agents, such as nivolumab and ipilimumab, are associated with AEs that differ in severity and duration than AEs caused by other therapeutic classes. Early recognition and management of AEs associated with immuno-oncology agents may mitigate severe toxicity. Management Algorithms have been developed to assist investigators in assessing and managing the following groups of AEs:

- Gastrointestinal
- Renal
- Pulmonary

- Hepatic
- Endocrinopathy
- Skin
- Neurological

The management algorithms recommended for use in CA2099ER are included in Appendix 5.

7.4.1.2 Dose Delay Criteria for Cabozantinib

Cabozantinib dosing should be delayed for the following:

- Urine protein/creatinine ratio (UPCR) > 2.0 *or* urine dipstick protein ≥ 3+ *or* urine protein > 2.0 g / 24 hours. Obtain 24 hour urine protein prior to next dosing visit.
- Grade 3 prolonged QTc interval (ie, QTcF interval > 500 msec on at least 2 out of 3 separate ECGs performed at least 3 minutes apart)
- Any Grade 2 or 3 drug-related venous thrombosis requiring anticoagulation, with the following exception:
  - Any recurrent or worsening venous thromboembolic event after restarting cabozantinib will require discontinuation
- Any other Grade 2 drug-related adverse event or grade 2 drug-related laboratory abnormality (e.g., AST, ALT, total bilirubin) that persists for more than 1 week or worsens despite supportive care management, with the following exception:
  - Any Grade 2 drug-related hemorrhage requires dose delay
  - Grade ≥ 2 arterial thromboembolic events, including cerebrovascular ischemia, cardiac ischemia/infarction, peripheral or visceral arterial ischemia, require discontinuation
- Sustained Grade 3 drug-related hypertension (systolic BP ≥ 160 mm Hg or diastolic BP ≥ 100 mm Hg). NOTE: Stopping or reducing the dose of cabozantinib is expected to cause a decrease in BP. The treating physician should monitor the participant for hypotension and adjust the number and dose of antihypertensive medications accordingly.
- Any other Grade 3 drug-related adverse event or laboratory abnormality, with the following exceptions:
  - Grade ≥ 3 drug-related hemorrhage requires discontinuation (Section 8.1.2).
  - Drug-related AST or ALT > 8 x ULN requires discontinuation (Section 8.1.2).
  - Concurrent AST or ALT > 3 x ULN and total bilirubin > 2 x ULN, consistent with potential drug-induced liver injury (see Section 9.2.7), requires discontinuation (Section 8.1.2)
  - Grade 3 drug-related amylase or lipase abnormalities that are not associated with symptoms or clinical manifestations of pancreatitis do not require dose delay. In such cases, more frequent monitoring (eg, weekly) of amylase and lipase is recommended. If amylase or lipase worsen to Grade 4 severity or the participant develops symptoms or clinical manifestations of pancreatitis, dosing should be delayed.
- Grade 4 drug-related amylase or lipase abnormalities require dose delay. Participants should be monitored for development of symptoms or clinical manifestations of pancreatitis.
- Grade 4 drug-related electrolyte abnormalities require dose delay. Electrolyte correction with supplementation/appropriate management should be promptly initiated.

- Grade 4 drug-related neutropenia, lymphopenia, leukopenia, anemia, or thrombocytopenia
- Any adverse event, laboratory abnormality, or intercurrent illness which, in the judgment of the investigator, warrants delaying cabozantinib dosing
- For participants scheduled for major surgery, including dental surgery which may impact bone healing, cabozantinib dosing should be delayed at least 28 days prior to scheduled surgery. The treating physician should use clinical judgment with regard to the risks and benefits of the planned surgical procedure if it is not possible to delay cabozantinib dosing for 28 days prior to the procedure. A delay of cabozantinib dosing of 5 to 7 days is recommended for healing for minor surgery.

As a general approach, all AEs related to cabozantinib should be managed with supportive care when possible at the earliest signs of toxicity. Calcium, magnesium, potassium, and phosphorus should be kept above the lower limits of the laboratory normal values. Please refer to the cabozantinib IB, Appendix 10, and Appendix 11 for additional information regarding dose modifications and AE management.[21]

7.4.1.3 Dose Delay Criteria for Sunitinib

Sunitinib dose delays should be based on instructions in the approved product label and should be considered for any severe or intolerable drug-related adverse events.

Within a cycle, missed doses of sunitinib should be skipped. Participants should never be dosed during the 2-week off period of each 6-week cycle, even if treatment delays occurred earlier in the cycle and therapy is ready to be resumed. If treatment is delayed past the end of the 6-week cycle, the start of the next cycle should be delayed until treatment with sunitinib resumes.

Prior to resuming therapy after a dose delay, refer to Section 7.4.2 for dose reduction recommendations and Section 8 for discontinuation criteria.

For this protocol, sunitinib dosing should be delayed for any of the following:

- Urine protein/creatinine ratio (UPCR) ≥ 2.0 *or* urine dipstick protein ≥ 3+ *or* urine protein ≥ 2.0 g / 24 hours. Obtain 24 hour urine protein prior to next dosing visit.
- Grade 3 prolonged QTc interval (ie, QTcF interval > 500 msec on at least 2 out of 3 separate ECGs performed at least 3 minutes apart)
- Any Grade 2 or 3 drug-related venous thrombosis requiring anticoagulation, with the following exception:
  - Any recurrent or worsening venous thromboembolic event after restarting sunitinib will require discontinuation
- Any other Grade 2 drug-related adverse event that persists or worsens despite supportive care management, with the following exception:
  - Any Grade 2 drug-related hemorrhage requires dose delay
  - Grade ≥ 2 arterial thromboembolic events, including cerebrovascular ischemia, cardiac ischemia/infarction, peripheral or visceral arterial ischemia, require discontinuation
- Grade 3 drug-related hypertension (systolic BP ≥ 160 mm Hg or diastolic BP ≥ 100 mm Hg). NOTE: Stopping or reducing the dose of sunitinib is expected to cause a decrease in BP. The

treating physician should monitor the participant for hypotension and adjust the number and dose of antihypertensive medications accordingly.

- Any other Grade 3 drug-related adverse event or laboratory abnormality, with the following exceptions:
  - Grade ≥ 3 drug-related hemorrhage requires discontinuation (Section 8.1.2).
  - Drug-related AST or ALT > 8 x ULN requires discontinuation (Section 8.1.2).
  - Concurrent AST or ALT > 3 x ULN and total bilirubin > 2 x ULN, consistent with potential drug-induced liver injury (see Section 9.2.7), requires discontinuation (Section 8.1.2)
  - Grade 3 drug-related amylase or lipase abnormalities that are not associated with symptoms or clinical manifestations of pancreatitis do not require dose delay. In such cases, more frequent monitoring (eg, weekly) of amylase and lipase is recommended. If amylase or lipase worsen to Grade 4 severity or the participant develops symptoms or clinical manifestations of pancreatitis, dosing should be delayed.
- Grade 4 drug-related amylase or lipase abnormalities require dose delay. Participants should be monitored for development of symptoms or clinical manifestations of pancreatitis.
- Grade 4 drug-related electrolyte abnormalities require dose delay. Electrolyte correction with supplementation/appropriate management should be promptly initiated.
- Grade 4 drug-related neutropenia, lymphopenia, leukopenia, anemia, or thrombocytopenia
- Any adverse event, laboratory abnormality, or intercurrent illness which, in the judgment of the investigator, warrants delaying sunitinib dosing
- For participants scheduled for major surgery, including dental surgery which may impact bone healing, sunitinib dosing should be delayed at least 28 days prior to scheduled surgery. The treating physician should use clinical judgment with regard to the risks and benefits of the planned surgical procedure if it is not possible to delay sunitinib dosing for 28 days prior to the procedure. A delay of sunitinib dosing of 5 to 7 days is recommended for healing for minor surgery.

As a general approach, all AEs related to sunitinib should be managed with supportive care when possible at the earliest signs of toxicity. Calcium, magnesium, potassium, and phosphorus should be kept above the lower limits of the laboratory normal values. Please refer to the sunitinib approved product label and Appendix 10 for additional information regarding AE monitoring and management.

7.4.2 Dose Reductions and Escalations

Dose reductions are permitted for cabozantinib and sunitinib but not for nivolumab and ipilimumab (see Table 7.4.2-1).

Table 7.4.2-1: Dose Level Modifications Table

Dose Level	**Cabozantinib (tablet dose expression)**	**Sunitinib (capsule dose expression)**	**Nivolumab (IV)**	**Ipilimumab (IV)**	**Nivolumab (IV)**
0 (starting dose)	40 mg daily	50 mg daily	3mg/kg	1 mg/kg	240 mg

Table 7.4.2-1: Dose Level Modifications Table

Dose Level	Cabozantinib (tablet dose expression)	Sunitinib (capsule dose expression)	Nivolumab (IV)	Ipilimumab (IV)	Nivolumab (IV)
-1	20 mg daily	37.5 mg daily	-	-	-
-2	20 mg every other day	25 mg daily	-	-	-

7.4.2.1 Dose Reduction and Escalation for Nivolumab and Ipilimumab

No dose reductions or dose escalations of nivolumab or ipilimumab are allowed.

7.4.2.2 Dose Reduction and Escalation for Cabozantinib

Dose reductions and dose escalations for adverse event management are allowed for cabozantinib (Table 7.4.2-1). Cabozantinib doses will not be re-escalated once reduced, unless a concomitant strong CYP3A4 inducer is started (see below). The only exception is participants who continue on cabozantinib alone after discontinuation of nivolumab alone (Arm A) or discontinuation of both nivolumab and ipilimumab (Arm B) who may re-escalate one dose level if the prior dose reduction was due to a toxicity felt by the investigator to have been mainly related to nivolumab and/or ipilimumab.

After toxicity requiring a dose delay has improved and meets the criteria to resume dosing (Section 7.4.3.2), participants who were receiving cabozantinib 40 mg daily prior to the delay will resume cabozantinib at 20 mg daily. Participants who were receiving cabozantinib 20 mg daily prior to the delay and require another dose delay will resume cabozantinib at 20 mg every other day. If more than 2 dose reductions are necessary (ie, reduction to less than 20 mg every other day), cabozantinib must be permanently discontinued (Section 8.1.2).

Participants who required a dose delay due to Grade 3 hypertension, which improved with antihypertensive medications, or any Grade 2 or 3 drug-related adverse event or asymptomatic laboratory abnormality that improved to Grade ≤ 1 within 7 days with supportive medical care may resume cabozantinib at the same dose or a reduced dose, at the discretion of the investigator.

Participants with asymptomatic Grade 2 drug-related AST, ALT or total bilirubin elevation, or Grade 3 drug-related lipase or amylase abnormalities that are not associated with symptoms or clinical manifestations of pancreatitis may reduce cabozantinib by one dose level, without delaying dosing, at the discretion of the investigator. See also section 7.4.1.2.

In Participants Who Start Taking a Concomitant Strong CYP3A4 Inhibitor: Reduce the daily cabozantinib dose by 20 mg (for example, from 40 mg to 20 mg daily). Resume the dose that was used prior to initiating the CYP3A4 inhibitor 2 to 3 days after discontinuation of the strong inhibitor (see Prescribing Information for cabozantinib).

7.4.2.3 Dose Reduction and Escalation for Sunitinib

Sunitinib Dose Reductions are permitted as per the approved product label for safety reasons or when a concomitant strong CYP3A4 inhibitor is needed (Appendix 9). Selection of an alternative

concomitant medication with minimal or no enzyme inhibition potential is recommended whenever possible.

After toxicity requiring a dose delay has improved and meets the criteria to resume dosing (Section 7.4.3.3), participants should resume sunitinib at one dose level reduction. Dose reductions should occur in 12.5 mg decrements. No more than 2 dose reductions are allowed. If more than 2 dose reductions are necessary (ie, reduction to less than 25 mg daily), the participant must be permanently discontinued (Section 8.1.3).

Participants who required a dose delay due to Grade 3 hypertension, which improved with antihypertensive medications, or any Grade 2 or 3 drug-related adverse event or asymptomatic laboratory abnormality that improved to Grade ≤ 1 within 7 days with supportive medical care may resume sunitinib at the same dose or a reduced dose, at the discretion of the investigator.

At the time a dose reduction is considered, also refer to Section 7.4.2.3 for dose delay recommendations and Section 8.1.3 for discontinuation criteria.

Sunitinib Dose Escalations are permitted as per the approved product label when a concomitant CYP3A4 inducer is needed (Appendix 9). Selection of an alternative concomitant medication with minimal or no enzyme induction potential is recommended whenever possible.

7.4.3 Criteria to Resume Treatment

7.4.3.1 Criteria to Resume Nivolumab and Ipilimumab Treatment

Delayed doses of nivolumab and/or ipilimumab should be administered as soon as the participant meets criteria to resume treatment. If a dose has been delayed, the participant should not wait until the next scheduled dosing date.

Participants may resume treatment with study drug when the drug-related AE(s) resolve to Grade ≤ 1 or baseline value, with the following exceptions:

- Participants may resume treatment in the presence of Grade 2 fatigue.
- Participants who have not experienced a Grade 3 drug-related skin AE may resume treatment in the presence of Grade 2 skin toxicity.
- For participants with Grade 2 AST, ALT, and/or total bilirubin abnormalities, dosing may resume when laboratory values return to baseline and management with corticosteroids, if needed, has been completed.
- Drug-related pulmonary toxicity, diarrhea, or colitis, must have resolved to baseline before treatment is resumed. Participants with persistent Grade 1 pneumonitis after completion of a steroid taper over at least 1 month may be eligible for retreatment if discussed with and approved by the BMS Medical Monitor or designee.
- Drug-related endocrinopathies adequately controlled with only physiologic hormone replacement may resume treatment.
- If treatment is delayed > 6 weeks, the participant must be permanently discontinued from study therapy, except as specified in Section 8.1.1.

For Arm B participants who delay nivolumab and ipilimumab dosing after Cycle 1 or 2, both nivolumab and ipilimumab must be resumed on the same day when the criteria to resume treatment are met.

For Arm B participants who delay nivolumab and ipilimumab dosing after Cycle 3 due to any drug-related AE meeting dose delay criteria that does not resolve within 14 days or requires treatment with systemic corticosteroids, it is acceptable to omit Cycle 4 if the investigator feels that ipilimumab was the main cause of the toxicity requiring dose delay. In this situation, when the participant meets criteria to resume nivolumab, the participant may proceed to Cycle 5 and begin nivolumab monotherapy maintenance.

7.4.3.2 Criteria to Resume Cabozantinib Treatment

Participants may resume treatment with cabozantinib when the drug-related AE(s) resolve to Grade ≤ 1 or baseline value, with the following exceptions:

- Participants may resume dosing in the presence of Grade 2 fatigue
- Participants who delayed dosing due to prolonged QTcF may resume dosing at one reduced dose level once QTcF returns to ≤ 500 msec
- Participants who delayed dosing due to UPCR > 2.0 or urine dipstick protein ≥ 3+ or urine protein > 2.0 g / 24 may resume dosing at one dose level reduction when UPCR is ≤ 2.0 or 24 hour urine protein ≤ 2.0 g /24
- Participants who delayed dosing due to Grade 3 hypertension may resume dosing at the same dose or at one dose level reduction, at the discretion of the investigator, when hypertension has improved to Grade ≤ 2
- Participants who delayed dosing due to Grade 4 lipase or amylase abnormalities may resume dosing upon resolution to Grade ≤ 2
- Participants who delayed dosing due to major surgery should not resume cabozantinib until complete wound healing has taken place. Following cabozantinib resumption, participants should be monitored for wound dehiscence, wound infections, and other signs of impaired wound healing.
- Participants who develop a pulmonary embolism and/or DVT should have study treatment interrupted until therapeutic anticoagulation is established. Treatment with cabozantinib may be resumed in subjects with pulmonary embolism or DVT if it is determined that the event is uncomplicated and that the subject is deriving clinical benefit from cabozantinib treatment and that anticoagulation does not place them at a significant risk that outweighs the benefit of resuming treatment per discretion of the investigator.
- Participants who delayed dosing due to ≤ Grade 2 hemorrhage may resume dosing if bleeding is under control (recovered to at least Grade 1 level) and has a low risk of recurrence
- If treatment is delayed > 6 weeks for any reason, the participant must be permanently discontinued from study therapy, except in cases where permission to resume treatment is granted by the BMS Medical Monitor or designee.

7.4.3.3 Criteria to Resume Sunitinib Treatment

Within a cycle, missed doses of sunitinib should be skipped and not replaced. Participants should never be dosed during the 2-week off period of each 6-week cycle, even if treatment delays occurred earlier in the cycle and therapy is ready to be resumed. If treatment is delayed past the end of the 6-week cycle, the start of the next cycle should be delayed until treatment with sunitinib resumes.

Participants may resume treatment with sunitinib when the drug-related AE(s) resolve to Grade ≤ 1 or baseline value, with the following exceptions:

- Participants may resume dosing in the presence of Grade 2 fatigue
- Participants who delayed dosing due to prolonged QTcF may resume dosing at one reduced dose level once QTcF returns to ≤ 500 msec
- Participants who delayed dosing due to UPCR ≥ 2.0 or urine dipstick protein ≥ 3+ may resume dosing at one dose level reduction when 24 hour urine protein < 2.0 g
- Participants who delayed dosing due to Grade 3 hypertension may resume dosing at the same dose or at one dose level reduction, at the discretion of the investigator, when hypertension has improved to Grade ≤ 2
- Participants who delayed dosing due to Grade 4 lipase or amylase abnormalities may resume dosing upon resolution to Grade ≤ 2
- Participants who develop a pulmonary embolism and/or DVT should have study treatment interrupted until therapeutic anticoagulation is established. Treatment with cabozantinib may be resumed in subjects with pulmonary embolism or DVT if it is determined that the event is uncomplicated and that the subject is deriving clinical benefit from cabozantinib treatment and that anticoagulation does not place them at a significant risk that outweighs the benefit of resuming treatment per discretion of the investigator.
- Participants who delayed dosing due to ≤ Grade 2 hemorrhage may resume dosing if bleeding is under control (recovered to at least Grade 1 level) and has a low risk of recurrence
- If treatment is delayed > 6 weeks for any reason, the participant must be permanently discontinued from study therapy, except in cases where permission to resume treatment is granted by the BMS Medical Monitor or designee.

7.4.4 Treatment of Nivolumab- or Ipilimumab-Related Infusion Reactions

Since nivolumab and ipilimumab contain only human immunoglobulin protein sequences, they are unlikely to be immunogenic and induce infusion or hypersensitivity reactions. However, if such a reaction were to occur, it might manifest with fever, chills, rigors, headache, rash, pruritus, arthralgias, hypotension, hypertension, bronchospasm, or other allergic-like reactions. Regardless of whether or not the event is attributed to the study drugs, all Grade 3 or 4 infusion reactions should be reported within 24 hours to the study BMS Medical Monitor or designee and reported as an SAE if it meets the criteria. Infusion reactions should be graded according to NCI CTCAE (Version 4) guidelines.

Treatment recommendations are provided below and may be modified based on local treatment standards and guidelines, as appropriate:

For Grade 1 symptoms: (mild reaction; infusion interruption not indicated; intervention not indicated):

- Remain at bedside and monitor participant until recovery from symptoms. The following prophylactic premedications are recommended for future infusions: diphenhydramine 50 mg (or equivalent) and/or acetaminophen/paracetamol 325 to 1000 mg at least 30 minutes before additional study drug administrations.

For Grade 2 symptoms: (moderate reaction required therapy or infusion interruption but responds promptly to symptomatic treatment (eg, antihistamines, non-steroidal anti-inflammatory drugs, narcotics, corticosteroids, bronchodilators, IV fluids); prophylactic medications indicated for ≤ 24 hours):

- Stop the study drug infusion, begin an IV infusion of normal saline, and treat the participant with diphenhydramine 50 mg IV (or equivalent) and/or acetaminophen/paracetamol 325 to 1000 mg; remain at bedside and monitor participant until resolution of symptoms. Corticosteroid and/or bronchodilator therapy may also be administered as appropriate. If the infusion is interrupted, then restart the infusion at 50% of the original infusion rate when symptoms resolve; if no further complications ensue after 30 minutes, the rate may be increased to 100% of the original infusion rate. Monitor participant closely. If symptoms recur, then no further study dug will be administered at that visit.
- For future infusions, the following prophylactic premedications are recommended: diphenhydramine 50 mg (or equivalent) and/or acetaminophen/paracetamol 325 to 1000 mg should be administered at least 30 minutes before study drug infusions. If necessary, corticosteroids (up to 25 mg of hydrocortisone or equivalent) may be used.

For Grade 3 or 4 symptoms: (severe reaction, Grade 3: prolonged [ie, not rapidly responsive to symptomatic medication and/or brief interruption of infusion]; recurrence of symptoms following initial improvement; hospitalization indicated for other clinical sequelae (eg, renal impairment, pulmonary infiltrates). Grade 4: Life-threatening; pressor or ventilatory support indicated):

- Immediately discontinue infusion of study drug. Begin an IV infusion of normal saline and treat the participant as follows: Recommend bronchodilators, epinephrine 0.2 to 1 mg of a 1:1000 solution for subcutaneous administration or 0.1 to 0.25 mg of a 1:10,000 solution injected slowly for IV administration, and/or diphenhydramine 50 mg IV with methylprednisolone 100 mg IV (or equivalent), as needed. Participant should be monitored until the Investigator is comfortable that the symptoms will not recur. Study drug will be permanently discontinued. Investigators should follow their institutional guidelines for the treatment of anaphylaxis. Remain at bedside and monitor participant until recovery of the symptoms.

In case of late-occurring hypersensitivity symptoms (eg, appearance of a localized or generalized pruritus within 1 week after treatment), symptomatic treatment may be given be given (eg, oral antihistamine or corticosteroids).

7.5 Preparation/Handling/Storage/Accountability

The investigational product should be stored in a secure area according to local regulations. It is the responsibility of the investigator to ensure that investigational product is only dispensed to

study Participants. The investigational product must be dispensed only from official study sites by authorized personnel according to local regulations.

The product storage manager should ensure that the study treatment is stored in accordance with the environmental conditions (temperature, light, and humidity) as determined by BMS. If concerns regarding the quality or appearance of the study treatment arise, the study treatment should not be dispensed and contact BMS immediately.

Study treatment not supplied by BMS will be stored in accordance with the package insert.

Investigational product documentation (whether supplied by BMS or not) must be maintained that includes all processes required to ensure drug is accurately administered. This includes documentation of drug storage, administration and, as applicable, storage temperatures, reconstitution, and use of required processes (eg, required diluents, administration sets).

- Further guidance and information for final disposition of unused study treatment are provided in Appendix 2.

7.5.1 Retained Samples for Bioavailability / Bioequivalence

Not applicable.

7.6 Treatment Compliance

Study treatment compliance will be periodically monitored by drug accountability (including review of dosing diary cards, as applicable). Drug accountability should be reviewed by the site study staff at each visit to confirm treatment compliance. Sites should discuss discrepancies with the participant at each on-treatment study visit.

7.7 Concomitant Therapy

7.7.1 Prohibited and/or Restricted Treatments

The following medications are prohibited during the study (unless utilized to treat a drug related adverse event):

- Immunosuppressive agents
- Immunosuppressive doses of systemic corticosteroids (except as stated in Section 7.7.3)
- Any concurrent anti-neoplastic therapy (ie, chemotherapy, hormonal therapy, immunotherapy, extensive, non-palliative radiation therapy, or standard or investigational agents)
- Chronic co-administration of cabozantinib with strong inducers of the CYP3A4 family (eg, phenytoin, carbamazepine, rifampin, rifabutin, rifapentine, phenobarbital, and St. John's Wort) may significantly decrease cabozantinib concentrations and must be avoided. Selection of alternate concomitant medications with no or minimal CYP3A4 enzyme induction potential is recommended. Caution must be used when discontinuing treatment with a strong CYP3A4 inducer in a subject who has been concurrently receiving a stable dose of cabozantinib, as this could significantly increase the exposure to cabozantinib.
- Any botanical preparation (eg herbal supplements or traditional Chinese medicines) intended to treat the disease under study or provide supportive care. Use of marijuana and its derivatives for treatment of symptoms related to cancer or cancer treatment are permitted if obtained by

medical prescription or if its use (even without a medical prescription) has been legalized locally.

- Any live/attenuated vaccine (eg varicella, zoster, yellow fever, rotavirus, oral polio and measles, mumps, rubella (MMR)) during treatment and until 100 days post last dose.

Supportive care for disease-related symptoms may be offered to all participants on the trial.

7.7.2 Other Restrictions and Precautions

Participants with a condition requiring systemic treatment with either corticosteroids (> 10 mg daily prednisone equivalent) or other immunosuppressive medications within 14 days of randomization are excluded. Inhaled or topical steroids, and adrenal replacement steroid doses > 10 mg daily prednisone equivalent, are permitted in the absence of active autoimmune disease.

7.7.2.1 Restrictions and Precautions for Cabozantinib

Cabozantinib is highly protein bound (99.9%) to human plasma proteins. Concomitant medications that are highly protein bound (eg, diazepam, furosemide, dicloxacillin, propanolol) should be used with caution. Because warfarin is a highly protein bound drug with a low therapeutic index, administration of warfarin should be avoided in participants receiving cabozantinib in Arms A and B.

Cabozantinib has been associated with a mild prolongation of the QTc interval. Caution should be used when treating participants on cabozantinib in Arms A and B with other drugs associated with QTc prolongation. Additional QTc monitoring is suggested for participants who are treated concomitantly with QTc prolonging drugs.

Low-dose aspirin for cardioprotection (as per local applicable guidelines) and low-dose low molecular weight heparins (LMWH) are permitted. Anticoagulation with therapeutic doses of LMWH is allowed in participants without known brain metastases who are on a stable dose of LMWH for at least 3 weeks before first dose of study treatment, and who have had no clinically significant hemorrhagic complications from the anticoagulation regimen or the tumor.

While on study, invasive dental procedures (eg, extractions, dental implants) should be avoided and conservative dental therapy, such as endodontic therapy (root canal treatment) would be preferred. After a surgical dental procedure, allow complete wound healing before resuming with cabozantinib treatment.

7.7.2.2 Restrictions and Precautions for Sunitinib

Sunitinib is metabolized by CYP3A4. Arm C participants who initiate CYP3A4 inhibitors/inducers after randomization should follow sunitinib dose modification instructions.

Sunitinib has been shown to prolong the QTc interval. Concomitant treatment with dysrhythmic drugs (ie, terfenadine, quinidine, procainamide, sotalol, probucol, bepridil, haloperidol, risperidone, and indapamide) is not recommended.

Medications taken within 4 weeks prior to study drug administration must be recorded on the CRF.

7.7.2.3 Imaging Restriction and Precautions

It is the local imaging facility's responsibility to determine, based on participant attributes (eg, allergy history, diabetic history and renal status), the appropriate imaging modality and contrast regimen for each participant. Imaging contraindications and contrast risks should be considered in this assessment. Participants with renal insufficiency should be assessed as to whether or not they should receive contrast and if so, what type and dose of contrast is appropriate. Specific to MRI, participants with severe renal insufficiency (ie, estimated glomerular filtration rate [eGFR] < 30 mL/min/1.73 m^2) are at increased risk of nephrogenic systemic fibrosis. MRI contrast should not be given to this participant population. In addition, participants are contraindicated from MRI if they have tattoos, metallic implants, pacemakers, etc.

The ultimate decision to perform MRI in an individual participant in this study rests with the site radiologist, the investigator and the standard set by the local Ethics Committee.

7.7.3 Permitted Therapy

Participants are permitted the use of topical, ocular, intra-articular, intranasal, and inhalational corticosteroids (with minimal systemic absorption). Adrenal replacement steroid doses > 10 mg daily prednisone are permitted. A brief (less than 3 weeks) course of corticosteroids for prophylaxis (eg, contrast dye allergy) or for treatment of non-autoimmune conditions (eg, delayed-type hypersensitivity reaction caused by a contact allergen) is permitted.

Concomitant medications are recorded at baseline and throughout the treatment phase of the study in the appropriate section of the CRF. All medications (prescriptions or over the counter medications) continued at the start of the study or started during the study and different from the study drug must be documented in the concomitant therapy section of the CRF.

7.8 Treatment After the End of the Study

At the conclusion of the study, participants who continue to demonstrate clinical benefit will be eligible to receive BMS supplied study treatment for the maximum treatment duration specified in protocol Section 7.1. Study treatment will be provided via an extension of the study, a rollover study requiring approval by responsible health authority and ethics committee or through another mechanism at the discretion of BMS.

BMS reserves the right to terminate access to BMS supplied study treatment if any of the following occur: a) the study is terminated due to safety concerns; b) the development of the nivolumab, ipilimumab, cabozantinib, or sunitinib is terminated for other reasons, including but not limited to lack of efficacy and/or not meeting the study objectives; c) the participant can obtain medication from a government sponsored or private health program. In all cases BMS will follow local regulations.

8 DISCONTINUATION CRITERIA

8.1 Discontinuation from Study Treatment

Participants MUST discontinue investigational product (and non-investigational product at the discretion of the investigator) for any of the following reasons:

- Participant's request to stop study treatment. Participants who request to discontinue study treatment will remain in the study and must continue to be followed for protocol specified follow-up procedures. The only exception to this is when a participant specifically withdraws consent for any further contact with him/her or persons previously authorized by participant to provide this information
- Any clinical adverse event (AE), laboratory abnormality or intercurrent illness which, in the opinion of the investigator, indicates that continued participation in the study is not in the best interest of the participant
- Termination of the study by Bristol-Myers Squibb (BMS)
- Loss of ability to freely provide consent through imprisonment or involuntarily incarceration for treatment of either a psychiatric or physical (eg, infectious disease) illness. (Note: Under specific circumstances, and only in countries where local regulations permit, a participant who has been imprisoned may be permitted to continue as a participant. Strict conditions apply and BMS approval is required.)
- Criteria listed in 8.1.1
- Disease progression of RCC or occurrence of a secondary malignancy which requires systemic therapy for treatment

Refer to the Schedule of Activities for data to be collected at the time of treatment discontinuation and follow-up and for any further evaluations that can be completed.

In the case of pregnancy, the investigator must immediately notify the BMS Medical Monitor/designee of this event. In most cases, the study treatment will be permanently discontinued in an appropriate manner (eg, dose tapering if necessary for participant safety). Please call the BMS Medical Monitor within 24 hours of awareness of the pregnancy. If the investigator determines a possible favorable benefit/risk ratio that warrants continuation of study treatment, a discussion between the investigator and the BMS Medical Monitor/designee must occur.

All participants who discontinue study treatment should comply with protocol specified follow-up procedures as outlined in Section 2. The only exception to this requirement is when a participant withdraws consent for all study procedures including post-treatment study follow-up or loses the ability to consent freely (ie, is imprisoned or involuntarily incarcerated for the treatment of either a psychiatric or physical illness).

If study treatment is discontinued prior to the participant's completion of the study, the reason for the discontinuation must be documented in the participant's medical records and entered on the appropriate case report form (CRF) page.

8.1.1 Nivolumab and Ipilimumab Dose Discontinuation (Arms A and B)

CA2099ER Global Revised Protocol 01

Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B. Participants previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinical planned events, per protocol.

The assessment for discontinuation of nivolumab, ipilimumab, and cabozantinib should be made separately for each study drug. Although there is overlap among the discontinuation criteria, if

discontinuation criteria are met for one study drug but not the other(s), it may be acceptable to continue treatment with the study drug(s) that are not felt to be related the toxicity, as specified below. If the investigator considers the toxicity to be related to all study drugs or is unable to determine which of the study drug(s) in Arm A or Arm B are the cause of toxicity, then all study drugs in the treatment regimen should be discontinued, and the recommendations for management of toxicity related to all study drugs should be promptly initiated.

Nivolumab and/or ipilimumab treatment should be permanently discontinued for any of the following:

- Any Grade 2 drug-related uveitis, eye pain or blurred vision that does not respond to topical therapy and does not improve to Grade 1 severity within the re-treatment period OR requires systemic treatment
- Any Grade 3 non-skin, drug-related adverse event lasting > 7 days, or recurs with the following exceptions for laboratory abnormalities, diarrhea, colitis, neurologic toxicity, drug-related uveitis, pneumonitis, bronchospasm, hypersensitivity reactions, infusion reactions, and endocrinopathies:
  - Grade 3 drug-related diarrhea, colitis, neurologic toxicity, uveitis, pneumonitis, bronchospasm, myocarditis, hypersensitivity reaction, or infusion reaction of any duration requires discontinuation. NOTE: The diagnosis of colitis should be supported by findings on colonoscopy whenever possible.
  - Grade 3 drug-related endocrinopathies, adequately controlled with only physiologic hormone replacement do not require discontinuation. Adrenal insufficiency requires discontinuation regardless of control with hormone replacement.
  - Grade 3 drug-related laboratory abnormalities do not require treatment discontinuation except:
    - Grade 3 drug-related thrombocytopenia > 7 days or associated with bleeding requires discontinuation
    - Any drug-related liver function test (LFT) abnormality that meets the following criteria require discontinuation:
      - Grade ≥ 3 drug-related AST, ALT or Total Bilirubin requires discontinuation*
      - Concurrent AST or ALT > 3 x ULN and total bilirubin > 2x ULN

        * In most cases of Grade 3 AST or ALT elevation, study drug(s) will be permanently discontinued. If the investigator determines a possible favorable benefit/risk ratio that warrants continuation of study drug(s), a discussion between the investigator and the BMS Medical Monitor/designee must occur.
- Any Grade 4 drug-related adverse event or laboratory abnormality (including but not limited to creatinine, AST, ALT, or Total Bilirubin), except for the following events which do not require discontinuation:
  - Grade 4 neutropenia ≤ 7 days
  - Grade 4 lymphopenia or leukopenia or asymptomatic amylase or lipase

  - Isolated Grade 4 electrolyte imbalances/abnormalities that are not associated with clinical sequelae and are corrected with supplementation/appropriate management within 72 hours of their onset
  - Grade 4 drug-related endocrinopathy adverse events, such as, hyper- or hypothyroidism, or glucose intolerance, which resolve or are adequately controlled with physiologic hormone replacement (corticosteroids, thyroid hormones) or glucose-controlling agents, respectively, may not require discontinuation after discussion with and approval from the BMS Medical Monitor or designee.
- Any event that leads to delay in dosing lasting > 6 weeks from the previous dose requires discontinuation, with the following exceptions:
  - Dosing delays to allow for prolonged steroid tapers to manage drug-related adverse events are allowed.
  - Dosing delays lasting > 6 weeks from the previous dose that occur for non-drug-related reasons may be allowed if approved by the BMS Medical Monitor or designee.

Prior to re-initiating treatment in a participant with a dosing delay lasting > 6 weeks, the BMS Medical Monitor or designee must be consulted. Tumor assessments should continue as per protocol even if dosing is delayed. Periodic study visits to assess safety and laboratory studies should also continue every 6 weeks or more frequently if clinically indicated during such dosing delays.

- Any adverse event, laboratory abnormality, or intercurrent illness which, in the judgment of the Investigator, presents a substantial clinical risk to the participant with continued nivolumab dosing.

Participants in Arm B who meet any of the discontinuation criteria above prior to Cycle 3 must discontinue both nivolumab and ipilimumab and may not receive nivolumab monotherapy maintenance.

Participants in Arm B who meet any of the discontinuation criteria above after Cycle 3 or Cycle 4 may be able to proceed to Cycle 5 (skipping Cycle 4 if needed) to begin nivolumab monotherapy maintenance if the toxicity is felt to be related mainly to ipilimumab, only after discussion with and approval by the BMS Medical Monitor or designee.

8.1.2 Cabozantinib Dose Discontinuation

Permanently discontinue cabozantinib for participants with any of the following:

- Any requirement for more than 2 cabozantinib dose reductions (ie, reduction to less than 20 mg every other day)
- Any Grade ≥ 2 drug-related arterial thromboembolic events, including cerebrovascular ischemia, cardiac ischemia/infarction, peripheral or visceral arterial ischemia
- Any Grade ≥ 3 drug-related hemorrhage
- Grade 4 hypertension or persistent Grade 3 hypertension despite optimal medical management and cabozantinib dose reduction
- Drug-related reversible posterior leukoencephalopathy syndrome

- Development of drug-related fistula or GI perforation
- Drug-related nephrotic syndrome
- Any drug-related liver function test (LFT) abnormality that meets the following criteria requires discontinuation:
  - AST or ALT > 8 x ULN.
  - Concurrent AST or ALT > 3 x ULN and total bilirubin >2 x ULN, consistent with potential drug-induced liver injury (see Section 9.2.7).
- Any Grade 4 drug-related adverse event or laboratory abnormality, with the following exceptions:
  - Grade 4 neutropenia ≤ 7 days
  - Grade 4 lymphopenia or leukopenia
  - Grade 4 amylase or lipase abnormalities that are not associated with symptoms or clinical manifestations of pancreatitis
  - Isolated Grade 4 electrolyte abnormalities that are not associated with clinical sequelae and are corrected with supplementation/appropriate management within 72 hours of their onset
- Any dosing interruption lasting > 6 weeks unless the BMS Medical Monitor or designee is consulted and agrees with the rationale for resuming therapy after a delay > 6 weeks. Note that tumor assessments should continue as per protocol even if dosing is interrupted.
- Any adverse event, laboratory abnormality, or intercurrent illness which, in the judgment of the Investigator, presents a substantial clinical risk to the participant with continued cabozantinib dosing.

8.1.3 Sunitinib Dose Discontinuation

Treatment with sunitinib should be permanently discontinued for any of the following:

- Any requirement for more than 2 sunitinib dose reductions (ie, reduction to less than 25 mg daily)
- Any Grade drug-related arterial thrombosis.
- Grade 4 drug-related hemorrhage or recurrent Grade 3 drug-related hemorrhage after dose reduction.
- Grade 4 drug-related symptomatic venous thrombosis.
- Grade 4 drug-related cardiac toxicity.
- Grade 4 hypertension or persistent Grade 3 hypertension despite optimal medical management and sunitinib dose reduction
- Any drug-related liver function test (LFT) abnormality that meets the following criteria requires discontinuation:
  - AST or ALT > 8 x ULN.
  - Concurrent AST or ALT > 3 x ULN and total bilirubin >2 x ULN, consistent with potential drug-induced liver injury (see Section 9.2.7).
- Any other Grade 4 drug-related adverse event or laboratory abnormality, with the following exceptions:

  - Grade 4 neutropenia ≤ 7 days
  - Grade 4 lymphopenia or leukopenia
  - Grade 4 amylase or lipase abnormalities that are not associated with symptoms or clinical manifestations of pancreatitis
  - Isolated Grade 4 electrolyte abnormalities that are not associated with clinical sequelae and are corrected with supplementation/appropriate management within 72 hours of their onset
- Any dosing interruption lasting > 6 weeks unless the BMS Medical Monitor or designee is consulted and agrees with the rationale for resuming therapy after a delay > 6 weeks. Note that tumor assessments should continue as per protocol even if dosing is interrupted.
- Any adverse event, laboratory abnormality, or intercurrent illness which, in the judgment of the Investigator, presents a substantial clinical risk to the participant with continued sunitinib dosing.

8.1.4 Treatment Beyond Disease Progression

Accumulating evidence indicates a minority of participants treated with immunotherapy or anti-angiogenic therapy may derive clinical benefit despite initial evidence of PD.[19,51,52]

Participants, regardless of study arm, will be permitted to continue treatment beyond initial RECIST 1.1 defined PD, assessed by the investigator, up to a maximum of 24 months from the date of first dose, as long as they meet the following criteria:

- Investigator-assessed clinical benefit
- Tolerance of study drug
- Stable performance status
- Treatment beyond progression will not delay an imminent intervention to prevent serious complications of disease progression (eg, CNS metastases)
- Participant provides written informed consent prior to receiving additional treatment with the study drug regimen. All other elements of the main consent including description of reasonably foreseeable risks or discomforts, or other alternative treatment options will still apply.

A radiographic assessment/ scan should be performed within 6 weeks of initial investigator-assessed progression to determine whether there has been a decrease in the tumor size or continued PD. The assessment of clinical benefit should be balanced by clinical judgment as to whether the participant is clinically deteriorating and unlikely to receive any benefit from continued treatment with nivolumab.

If the investigator feels that the participant continues to achieve clinical benefit by continuing treatment, the participant should remain on the trial and continue to receive monitoring according to the Schedule of Activities Schedule in Section 2. Treatment may be continued beyond investigator-assessed progression if the investigator confirms that the participant meets the criteria specified in Section 8.1.4.

For the participants who continue study therapy beyond progression, further progression is defined as an additional 10% increase in tumor burden with a minimum 5 mm absolute increase from time of initial PD. This includes an increase in the sum of diameters of all target lesions and/ or the

diameters of new measurable lesions compared to the time of initial PD. It is recommended that study treatment should be discontinued permanently upon documentation of further progression.

New lesions are considered measureable at the time of initial progression if the longest diameter is at least 10 mm (except for pathological lymph nodes which must have a short axis of at least15 mm). Any new lesion considered non-measureable at the time of initial progression may become measureable and therefore included in the tumor burden if the longest diameter increases to at least 10 mm (except for pathological lymph nodes which must have a short axis of at least 15 mm). In situations where the relative increase in total tumor burden by 10% is solely due to inclusion of new lesions which become measurable, these new lesions must demonstrate an absolute increase of at least 5 mm.

8.1.5 Post-Study Treatment Follow-up

In this study, overall survival is a key endpoint of the study. Post-study treatment follow-up is of critical importance and is essential to preserving participant safety and the integrity of the study. Participants who discontinue study treatment must continue to be followed for collection of outcome and/or survival follow-up data as required and in line with Section 5 until death or the conclusion of the study.

BMS may request that survival data be collected on all treated/randomized participants outside of the protocol defined window (See Section 2). At the time of this request, each participant will be contacted to determine their survival status unless the participant has withdrawn consent for all contacts or is lost to follow-up.

8.2 Discontinuation from the Study

Participants who request to discontinue study treatment will remain in the study and must continue to be followed for protocol specified follow-up procedures. The only exception to this is when a participant specifically withdraws consent for any further contact with him/her or persons previously authorized by participant to provide this information.

- Participants should notify the investigator of the decision to withdraw consent from future follow-up **in writing**, whenever possible.
- The withdrawal of consent should be explained in detail in the medical records by the investigator, as to whether the withdrawal is from further treatment with study treatment only or also from study procedures and/or post treatment study follow-up, and entered on the appropriate CRF page.
- In the event that vital status (whether the participant is alive or dead) is being measured, publicly available information should be used to determine vital status only as appropriately directed in accordance with local law.
- If the participant withdraws consent for disclosure of future information, the sponsor may retain and continue to use any data collected before such a withdrawal of consent.

8.3 Lost to Follow-Up

- All reasonable efforts must be made to locate participants to determine and report their ongoing status. This includes follow-up with persons authorized by the participant.

- Lost to follow-up is defined by the inability to reach the participant after a minimum of **three** documented phone calls, faxes, or emails as well as lack of response by participant to one registered mail letter. All attempts should be documented in the participant's medical records.
- If it is determined that the participant has died, the site will use permissible local methods to obtain date and cause of death.
- If investigator's use of third-party representative to assist in the follow-up portion of the study has been included in the participant's informed consent, then the investigator may use a Sponsor retained third-party representative to assist site staff with obtaining participant's contact information or other public vital status data necessary to complete the follow-up portion of the study.
- The site staff and representative will consult publicly available sources, such as public health registries and databases, in order to obtain updated contact information.
- If after all attempts, the participant remains lost to follow-up, then the last known alive date as determined by the investigator should be reported and documented in the participant's medical records.

9 STUDY ASSESSMENTS AND PROCEDURES

CA2099ER Global Revised Protocol 01

Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B. Participants previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinical planned events, per protocol.

- Study procedures and timing are summarized in the Schedule of Activities.
- Protocol waivers or exemptions are not allowed.
- All immediate safety concerns must be discussed with the Sponsor immediately upon occurrence or awareness to determine if the participant should continue or discontinue treatment.
- Adherence to the study design requirements, including those specified in the Schedule of Activities, is essential and required for study conduct.
- All screening evaluations must be completed and reviewed to confirm that potential participants meet all eligibility criteria before randomization. The investigator will maintain a screening log to record details of all participants screened and to confirm eligibility or record reasons for screening failure, as applicable.
- Procedures conducted as part of the participant's routine clinical management (eg, blood count) and obtained before signing of informed consent may be utilized for screening or baseline purposes provided the procedure meets the protocol-defined criteria and has been performed within the timeframe defined in the Schedule of Activities.

Additional measures, including non-study required laboratory tests, should be performed as clinically indicated or to comply with local regulations. Laboratory toxicities (eg, suspected drug induced liver enzyme evaluations) will be monitored during the follow-up phase via on site/local labs until all study drug related toxicities resolve, return to baseline, or are deemed irreversible.

If a participant shows pulmonary-related signs (hypoxia, fever) or symptoms (eg, dyspnea, cough, fever) consistent with possible pulmonary adverse events, the participant should be immediately

evaluated to rule out pulmonary toxicity, according to the suspected pulmonary toxicity management algorithm in the BMS-936558 (nivolumab) Investigator Brochure.

Some of the assessments referred to in this section may not be captured as data in the eCRF. They are intended to be used as safety monitoring by the treating physician. Additional testing or assessments may be performed as clinically necessary or where required by institutional or local regulations.

9.1 Efficacy Assessments

Study evaluations will take place in accordance with the Schedule of Activities in Section 2.

Images will be submitted to an imaging core lab. Sites should be trained prior to scanning the first study participant. Image acquisition guidelines and submission process will be outlined in the CA2099ER Imaging Manual to be provided by the core lab.

9.1.1 Methods of Measurements

The following imaging assessments should be performed at pre-specified intervals: CT of the chest, CT or MRI of the abdomen, pelvis, and other known sites of disease.

- CT scans should be acquired with slice thickness of 5 mm or less with no intervening gap (continuous)
- Should a participant have a contraindication for CT IV contrast, a non-contrast CT of the chest and a contrast enhanced MRI of the abdomen and pelvis and other sites of disease may be obtained. MRIs should be acquired with slice thickness of 5 mm or less with no gap (continuous).
- Every attempt should be made to image each participant using an identical acquisition protocol on the same scanner for all imaging time points
- PET alone will not be considered for the disease assessment. Complementary CT and/or MRI or biopsy must be performed in such cases.

  Note: Use of CT component of a PET/CT scanner: Combined modality scanning, such as with FDG-PET/CT, is increasingly used in clinical care and is a modality/technology that is in rapid evolution; therefore, the recommendations outlined here may change rather quickly with time. At present, low-dose or attenuation correction CT portions of a combined FDG-PET/CT are of limited use in anatomically-based efficacy assessments, and it is, therefore, suggested that they should not be substituted for dedicated diagnostic contrast enhanced CT scans for anatomically-based RECIST (Appendix 8) measurements. However, if a site can document that the CT performed as part of a FDG-PET/CT is of identical diagnostic quality to a diagnostic CT (with IV and oral contrast), then the CT portion of the FDG-PET/CT can be used for RECIST 1.1 measurements. Note, however, that the FDG-PET portion of the CT introduces additional data which may bias an investigator if it is not routinely or serially performed.

9.1.2 Imaging and Clinical Assessment

Images will be submitted to an imaging core lab. Sites should be trained prior to scanning the first study participant. Image acquisition guidelines and submission process will be outlined in the CA2099ER Imaging Manual to be provided by the core lab.

Baseline imaging, including CT/MRI of the chest, abdomen, pelvis, and all known sites of disease performed within 28 days prior to randomization, should be submitted to the imaging core lab. Baseline brain MRI (preferred) or CT scan should also be performed within 28 days prior to randomization and submitted to the imaging core lab. Participants who are found to have untreated brain metastases on the baseline brain scan may not be randomized.

9.1.2.1 Investigator Assessment of Progression

The same method of assessment used at Screening should be used for on-study time points. Brain MRI or CT scans during on-study time points and the follow-up phase are only required in participants with a history of CNS metastases prior to randomization or if clinically indicated for new signs or symptoms that suggest new or worsening CNS metastases.

Post-baseline tumor assessments will be performed at the time points described below until progression assessed by the investigator **and** confirmed by BICR, death, or withdrawal from the study, whichever occurs first.

- First tumor assessment post-baseline should be performed at Week 12 (± 7 days) following randomization. Use same imaging method as was used at screening/baseline.
- Subsequent tumor assessments should occur at every 6 weeks until Week 60. Allowable window for assessments is ± 7 days until Week 60. After Week 60, tumor assessments should occur every 12 weeks (± 14 days) until radiographic progression, assessed by the investigator **and** confirmed by the BICR.
- Additional imaging of potential disease sites should be performed whenever disease progression or a secondary malignancy is suspected. In participants with no history of brain lesions prior to randomization, brain MRI or CT on-study treatment should be obtained if clinically indicated. Bone imaging during on-study treatment and follow-up periods should be obtained if clinically indicated.
- Tumor assessments are to continue for all randomized participants according to the protocol schedule until radiographic progression has been assessed by the investigator **and** confirmed by the BICR, regardless of whether study drug dosing is delayed, reduced, or discontinued.

The investigator (in consultation with the local radiologist as needed) will complete tumor assessments using RECIST (Response Evaluation Criteria in Solid Tumors) 1.1 criteria on all imaging time points specified in the protocol. Any additional imaging that may demonstrate tumor response or progression, including scans performed at unscheduled time points and/or at an outside institution, should also be collected for the investigator to complete RECIST 1.1 tumor assessments on these images and to submit them for BICR review.

All study treatment decisions will be based on the investigator's assessment of tumor images and not on the BICR assessment.

9.1.2.2 BICR Assessment of Progression

Sites should submit all scans to a BICR on a rolling basis, preferably within 7 days of scan acquisition, throughout the duration of the study. BICR will review scans on a rolling basis and remain blinded to treatment arm and investigator assessment of submitted scans. When progression per RECIST 1.1 criteria is assessed by the investigator, the site will inform the imaging core lab, so that the BICR assessment of progression can be performed. The BICR review will be completed and the results provided to the site within approximately 14 days of receipt of the scans, provided there are no pending imaging queries to the site.

Participants whose progression is not confirmed by the BICR will be required to continue tumor assessments (if clinically feasible) according to the protocol-specified schedule or sooner if clinically indicated until the BICR confirms progression on a subsequent tumor assessment. Also, if participants discontinue treatment without radiographic progression, tumor assessments will continue according to the protocol specified schedule, as noted in Section 2, until progression has been confirmed by BICR.

All study treatment decisions will be based on the investigator's assessment of tumor images and not on the BICR assessment. The BICR assessment of progression is only relevant for determining when tumor assessments for a given participant are no longer required to be submitted to the imaging vendor.

9.1.3 Imaging Restriction and Precautions

Table 9.1.3-1 provides a summary of the alternative methods, acceptable per protocol, in the event of contraindications for use of IV and oral contrast, and or/MRI.

Table 9.1.3-1: Acceptable Imaging Assessment Methods for Different Anatomic Regions

Anatomic Region	Preferred Method	Alternative Methods
Chest, abdomen, and pelvis Note: Scan must cover lung apices to diaphragm, diaphragm through entire liver, and to below the pubic symphysis	CT with IV contrast	**For chest:** • CT without contrast can be used only if the participant has a clinical contraindication for iodine-based IV contrast (eg, hypersensitivity, renal insufficiency) **For abdomen and pelvis:** • MRI with gadolinium-based IV contrast is the first alternative method if the participant has a clinical contraindication for iodine-based IV contrast • CT without contrast can be used as the second alternative method only if the participant has a clinical contraindication for both contrast-enhanced CT and MRI.
Brain	MRI with IV contrast	• CT with IV contrast is the first alternative method if IV gadolinium is clinically contraindicated.

Table 9.1.3-1: Acceptable Imaging Assessment Methods for Different Anatomic Regions

Anatomic Region	Preferred Method	Alternative Methods
		• MRI without contrast can be used as a second alternative method if a participant has clinical contraindications for both contrast-enhanced CT and MRI
Bone	Bone scintigraphy	PET (18F-fluoride NaF or FDG) and 99m Technetium SPECT

Notes:

1. CT scans must be performed with slices thickness of ≤ 5 mm are required. The reconstruction interval should be equal to slice thickness to avoid gap.
2. The same modality for a given anatomical coverage and the same scanning procedure (most importantly: reconstruction slice thickness, anatomic coverage, use of IV contrast) should be consistent between baseline and all subsequent follow-up scanning. If possible, the same scanner or an equivalent scanners should be used throughout the study.
3. For abdomen and pelvis CT scans, oral contrast is recommended as per institutional standards.
4. MRI should include T1 and T2-weighted sequences with T1-weighted at both pre- and post-contrast.
5. If bone scan shows hotspots indicative of metastases, further investigation with X-ray, CT, or MRI is warranted.
6. All scans generated should be exportable in electronic format (DICOM) to enable secure and rapid electronic transmission to the designated central imaging laboratory.

The use of gadolinium-based contrast agents in participants with acute or chronic renal insufficiency, with a glomerular filtration rate (GFR) less than 30 mL per minute per 1.73m^2 or with any acute renal failure caused by hepatorenal syndrome or perioperative liver transplantation, is not recommended.

If gadolinium is contraindicated, proceed without contrast but reason for not using contrast must be documented.

9.1.4 Outcomes Research Assessments

Participant-reported outcomes will be captured through the use of 2 validated self-reported questionnaires: the National Comprehensive Cancer Network (NCCN) Functional Assessment of Cancer Therapy - Kidney Symptom Index (FKSI-19), and the EuroQoL Group's EQ-5D-3L.

9.1.4.1 FKSI-19

The NCCN FKSI-19 [53 54] is a 19-item scale that measures tumor specific HrQoL in kidney cancer participants. The FKSI-19 uses 5 Likert-type response categories that range from "not at all" to "very much." Participants are asked to circle the response category that best characterizes their response over the last 7 days on 19 items that include symptoms such as lack of energy, fatigue, appetite, coughing, shortness of breath, pain, nausea, and ability to work. The instrument yields a total score and three subscale scores: Disease Related Symptoms (DRS), Treatment Side Effects (TSE), and Functional Well Being (FWB). A higher score indicates fewer symptoms.

9.1.4.2 EQ-5D-3L

The 3-level version of the EQ-5D (EQ-5D-3L)[55] will be used to assess treatment effects on perceived health status and to generate utility data for health economic evaluations. The EQ-5D-3L is a generic multi-attribute health-state classification system by which health is described in 5 dimensions (ie, mobility, self-care, usual activities, pain/discomfort, and anxiety) and depression. Each dimension is evaluated using 3 levels: no problems, some problems, and severe problems.

Responses to these 5 dimensions are converted into 1 of 243 unique EQ-5D health state descriptions, which range between no problems on all 5 dimensions [11111] to severe/extreme problems on all 5 dimensions [33333]. Using appropriate country-specific value weighting algorithms, a respondent's self-described health state can be converted into a utility index representing the societal desirability of his/her own health. In addition, the EQ-5D includes a VAS allowing a respondent to rate his/her health on a scale 0–100, with 0 being the worst health state and 100 being the best health state imaginable.

9.2 Adverse Events

The definitions of an AE or serious adverse event (SAE) can be found in Appendix 3.

AEs will be reported by the participant (or, when appropriate, by a caregiver, surrogate, or the participant's legally authorized representative).

The investigator and any designees are responsible for detecting, documenting, and reporting events that meet the definition of an AE or SAE and remain responsible for following up AEs that are serious, considered related to the study treatment or the study, or that caused the participant to discontinue before completing the study.

Contacts for SAE reporting specified in Appendix 3

Immune-mediated adverse events (IMAEs, imAEs) are AEs consistent with an immune-mediated mechanism or immune-mediated component for which non-inflammatory etiologies (eg, infection or tumor progression) have been ruled out. IMAEs can include events with an alternate etiology which were exacerbated by the induction of autoimmunity. Information supporting the assessment will be collected on the participant's case report form.

9.2.1 Time Period and Frequency for Collecting AE and SAE Information

The collection of nonserious AE information should begin at initiation of study treatment until the timepoints specified in the Schedule of Activities (Section 2). Nonserious AE information should also be collected from the start of a placebo lead-in period or other observational period intended to establish a baseline status for the participants.

Sections 5.6.1 and 5.6.2 in the Investigator Brochures (IBs) for Nivolumab[28] and Ipilimumab[29] represent the Reference Safety Information (also Appendix K in the IB for cabozantinib[21]) to determine expectedness of serious adverse events for expedited reporting. Following the participant's written consent to participate in the study, all SAEs, whether related or not related to study drug, must be collected, including those thought to be associated with protocol-specified procedures.

All SAEs must be collected that occur during the screening period and within 100 days of the last dose of study treatment. For participants randomized to treatment and never treated with study drug, SAEs should be collected for 30 days from the date of randomization. If applicable, SAEs must be collected that relate to any later protocol-specified procedure (eg, a follow-up skin biopsy).

The investigator must report any SAE that occurs after these time periods and that is believed to be related to study drug or protocol-specified procedure.

- Medical occurrences that begin before the start of study treatment but after obtaining informed consent will be recorded on the appropriate section of the eCRF section.
- All SAEs will be recorded and reported to Sponsor or designee within 24 hours, as indicated in Appendix 3.
- The investigator will submit any updated SAE data to the sponsor within 24 hours of this being available.

Investigators are not obligated to actively seek AEs or SAEs in former study participants. However, if the investigator learns of any SAE, including a death, at any time after a participant has been discharged from the study, and he/she considers the event reasonably related to the study treatment or study participation, the investigator must promptly notify the sponsor.

The method of evaluating, and assessing causality of AEs and SAEs and the procedures for completing and reporting/transmitting SAE reports are provided in Appendix 3.

9.2.2 Method of Detecting AEs and SAEs

Adverse events can be spontaneously reported or elicited during open-ended questioning, examination, or evaluation of a participant. Care should be taken not to introduce bias when collecting AE and/or SAEs. Inquiry about specific AEs should be guided by clinical judgement in the context of known adverse events, when appropriate for the program or protocol.

9.2.3 Follow-up of AEs and SAEs

- Nonserious AEs should be followed to resolution or stabilization, or reported as SAEs if they become serious (see Appendix 3).
- Follow-up is also required for nonserious AEs that cause interruption or discontinuation of study treatment and for those present at the end of study treatment as appropriate.
- All identified nonserious AEs must be recorded and described on the nonserious AE page of the CRF (paper or electronic). Completion of supplemental CRFs may be requested for AEs and/or laboratory abnormalities that are reported/identified during the course of the study.

After the initial AE/SAE report, the investigator is required to proactively follow each participant at subsequent visits/contacts. All SAEs, and non-serious AEs of special interest (as defined in Section 9.2 will be followed until resolution, until the condition stabilizes, until the event is otherwise explained, or until the participant is lost to follow-up (as defined in Section 8.3).

Further information on follow-up procedures is given in Appendix 3.

9.2.4 Regulatory Reporting Requirements for SAEs

- Prompt notification by the investigator to the Sponsor of SAEs is essential so that legal obligations and ethical responsibilities towards the safety of participants and the safety of a product under clinical investigation are met.
- An investigator who receives an investigator safety report describing SAEs or other specific safety information (eg, summary or listing of SAEs) from the Sponsor will file it along with the Investigator Brochure and will notify the IRB/IEC, if appropriate according to local requirements.

Sponsor or designee will be reporting adverse events to regulatory authorities and ethics committees according to local applicable laws including European Directive 2001/20/EC and FDA Code of Federal Regulations 21 CFR Parts 312 and 320. A SUSAR (Suspected, Unexpected Serious Adverse Reaction) is a subset of SAEs and will be reported to the appropriate regulatory authorities and investigators following local and global guidelines and requirements.

9.2.5 Pregnancy

If, following initiation of the study treatment, it is subsequently discovered that a participant is pregnant or may have been pregnant at the time of study exposure, including during at least 5 months (5 half-lives + 30 days) after product administration for WOCBP, 7 months (5 half lives + 90 days) after product administration for male participants with partners who are WOCBP, the investigator must immediately notify the BMS Medical Monitor/designee of this event and complete and forward a Pregnancy Surveillance Form to BMS Designee within 24 hours of awareness of the event and in accordance with SAE reporting procedures described in Appendix 3.

In most cases, the study treatment will be permanently discontinued in an appropriate manner (eg, dose tapering if necessary for participant safety). Please call the BMS Medical Monitor within 24 hours of awareness of the pregnancy.

Protocol-required procedures for study discontinuation and follow-up must be performed on the participant.

Follow-up information regarding the course of the pregnancy, including perinatal and neonatal outcome and, where applicable, offspring information must be reported on the Pregnancy Surveillance Form.

Any pregnancy that occurs in a female partner of a male study participant should be reported to Sponsor or designee. In order for Sponsor or designee to collect any pregnancy surveillance information from the female partner, the female partner must sign an informed consent form for disclosure of this information. Information on this pregnancy will be collected on the Pregnancy Surveillance Form.

9.2.6 Laboratory Test Result Abnormalities

The following laboratory test result abnormalities should be captured on the nonserious AE CRF page or SAE Report Form electronic, as appropriate. Paper forms are only intended as a back-up option when the electronic system is not functioning.

- Any laboratory test result that is clinically significant or meets the definition of an SAE
- Any laboratory test result abnormality that required the participant to have study treatment discontinued or interrupted
- Any laboratory test result abnormality that required the participant to receive specific corrective therapy

It is expected that wherever possible, the clinical rather than laboratory term would be used by the reporting investigator (eg, anemia versus low hemoglobin value).

9.2.7 Potential Drug Induced Liver Injury (DILI)

Wherever possible, timely confirmation of initial liver-related laboratory abnormalities should occur prior to the reporting of a potential DILI event. All occurrences of potential DILIs, meeting the defined criteria, must be reported as SAEs (see Section 9.2 and Appendix 3 for reporting details).

Potential drug induced liver injury is defined as:

- AT (ALT or AST) elevation > 3 times upper limit of normal (ULN)

  AND
- Total bilirubin > 2 times ULN, without initial findings of cholestasis (elevated serum alkaline phosphatase),

  AND
- No other immediately apparent possible causes of AT elevation and hyperbilirubinemia, including, but not limited to, viral hepatitis, pre-existing chronic or acute liver disease, or the administration of other drug(s) known to be hepatotoxic.

9.2.8 Other Safety Considerations

Any significant worsening noted during interim or final physical examinations, electrocardiogram, x-ray filming, any other potential safety assessment required or not required by protocol should also be recorded as a nonserious or serious AE, as appropriate, and reported accordingly.

9.3 Overdose

An overdose is defined as the accidental or intentional administration of any dose of a product that is considered both excessive and medically important. All occurrences of overdose must be reported as SAEs (see Section 9.2).

9.4 Safety

Planned time points for all safety assessments are listed in the Schedule of Activities (see Section 2).

9.4.1 Clinical Safety Laboratory Assessments

Investigators must document their review of each laboratory safety report.

Hematology (CBC)
Hemoglobin
Hematocrit

Total leukocyte count, including differential	
Platelet count	
Prothrombin time (PT)/ International normalized ratio (INR)	
Partial thromboplastin time (PTT)	
Chemistry	
Aspartate aminotransferase (AST) Alanine aminotransferase (ALT) Total bilirubin Alkaline phosphatase Lactate dehydrogenase (MLR) Creatinine Blood Urea Nitrogen (BUN) Glucose Albumin	Sodium (Na) Potassium (K) Chloride (Cl) Calcium (Ca) Corrected calcium (Screening only) Phosphorus (P) Magnesium (Mg) Amylase Lipase
Urinalysis	
Creatinine Protein Urine protein/creatinine ratio (UPCR)	
Serology	
Serum for hepatitis C antibody, HCV RNA, hepatitis B surface antigen (screening only). HIV, if mandated locally. (Sites in Germany, see Appendix 12)	
Other Analyses	
Thyroid stimulating hormone (TSH) with free thyroxine (fT3) and free triiodothyronine (fT4) (screening only); TSH with reflexive fT3 and fT4 during study and follow up	
Pregnancy test (WOCBP only: screening and during study)	
Follicle stimulating hormone (FSH) (screening only for women under 55 years old to confirm menopause as needed)	

9.4.2 Imaging Safety Assessment

Any incidental findings of potential clinical relevance that are not directly associated with the objectives of the protocol should be evaluated and handled by the Study Investigator as per standard medical/clinical judgment.

9.5 Pharmacokinetics and Immunogenicity

CA2099ER Global Revised Protocol 01

Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B. Participants previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinical planned events, per protocol.

Samples for nivolumab, ipilimumab, and cabozantinib PK and nivolumab and ipilimumab immunogenicity assessments will be collected for participants in Arm A (doublet) receiving nivolumab combined with cabozantinib and in Arm B (triplet) receiving nivolumab and ipilimumab combined with cabozantinib as described in Table 9.5-1 and Table 9.5-2, respectively.

Ipilimumab samples will be collected and only analyzed if needed. Treatment assignments will be released to the bioanalytical laboratory in order to minimize unnecessary analysis of samples. All timepoints for nivolumab/ipilimumab sampling are relative to the start of study drug administration for nivolumab in Arms A and B. All timepoints for cabozantinib PK sampling are relative to the start of study drug administration for nivolumab in Arms A and B and should be drawn approximately at least 8 hours after the previous evening cabozantinib dose.

All on-treatment timepoints are intended to align with days on which study drug is administered. If it is known that a dose is going to be delayed, then the predose sample for nivolumab/ipilimumab and cabozantinib, if appropriate, should be collected just prior to the delayed dose. However, if a predose sample for nivolumab/ipilimumab is collected but the dose is subsequently delayed, an additional predose sample should not be collected. Further details of sample collection, processing, and shipment will be provided in the laboratory/procedure manual.

Blood samples should be drawn from a site other than the infusion site (ie, contralateral arm) on days of infusion. If the infusion was interrupted, the interruption details will also be documented on the CRF. Further details of pharmacokinetic sample collection and processing will be provided to the site in the laboratory/procedure manual.

Samples collected from participants for immunogenicity analyses of nivolumab be evaluated for the development of Anti-Drug Antibody (ADA) by validated immunoassays. Samples may also be analyzed for neutralizing ADA.

Serum concentration analyses for nivolumab and/or ipilimumab will be performed by validated immunoassay bioanalytical method(s) for nivolumab and ipilimumab. Plasma concentration analyses for cabozantinib will be performed by validated liquid chromatography tandem-mass spectrometry (LC-MS-MS).

In addition, selected serum samples may be analyzed by an exploratory method that measures nivolumab and ipilimumab, or detect anti-drug antibodies for technology exploration purposes; exploratory results will not be reported. The corresponding serum samples designated for either PK, immunogenicity or biomarker assessments may also be used for any of those analyses, if required (eg, insufficient sample volume to complete testing or to follow up on suspected immunogenicity related AE).

Table 9.5-1: Pharmacokinetic and Immunogenicity Sampling Schedule for Arm A (Doublet)

Study Day[a] (1 Cycle = 2 weeks)	Event (Relative to Nivolumab Dosing) Hour	Time (Relative to Start of Nivolumab Infusion) Hours:Min	Pharmacokinetic Blood Sample for Nivolumab for Arm A	Immunogenicity Blood Sample for Nivolumab for Arm A	Pharmacokinetic Blood Sample for Cabozantinib for Arm A
C1/D1	Predose nivo[b]	00:00	X	X	X[c]
C1/D1	EOI-nivo[d]	00:30	X		
C3/D1	Predose nivo	00:00			X[e]
C4/D1	Predose nivo	00:00	X	X	X[e]
C7/D1	Predose nivo	00:00	X	X	X[e]
C15/D1	Predose nivo	00:00	X	X	
C23/D1	Predose nivo	00:00	X	X	
Every 16 weeks from C23 to 2 years	Predose nivo	00:00	X	X	
Follow-up Samples - Approximately 30 and 100 days from the discontinuation of study drug	NA	NA	X	X	

Abbreviations: C=cycle; D=day; EOI=end of infusion; NA=not applicable; nivo=nivolumab

[a] If a participant discontinues study drug treatment during the sampling period, they will move to sampling at the follow-up visits.

[b] Predose: All predose samples for nivolumab should be taken prior to the start of nivolumab infusion.

[c] Though cabozantinib will be dosed in the evening on Day 1 of Cycle 1, the cabozantinib predose sample on Day 1 can be drawn at the same time nivolumab predose sample is drawn.

[d] EOI-nivo : End of Infusion samples for nivolumab should be collected immediately (preferably within 2-5 minutes) prior to the end of infusion. If the end of infusion is delayed, the collection of the EOI samples should be delayed accordingly. EOI samples may not be collected from the same IV access as drug was administered, refer to the laboratory manual for additional instructions.

[e] Cabozantinib is preferably dosed at bedtime. A PK sample for cabozantinib will be drawn at the same time when predose PK samples for nivolumab are drawn as long at the time of the draw for cabozantinib is approximately at least 8 hours after the previous evening dose of cabozantinib.

Table 9.5-2: Pharmacokinetic (PK) and Immunogenicity Sampling Schedule for Arm B (Triplet)

Study Day[a] (C1 to C4 Cycles = 3 weeks then C5 onward Cycles =2 weeks)	Event (Relative to Nivolumab Dosing) Hour	Time (Relative to Start of Nivolumab Infusion) Hours:Min	PK Blood Sample for Nivolumab for Arm B	Immunogenicity Blood Sample for Nivolumab for Arm B	PK Blood Sample for Ipilimumab for Arm B[b]	Immunogenicity Blood Sample for Ipilimumab for Arm B[b]	PK Blood Sample for Cabozantinib for Arm B
C1/D1	Predose nivo[c]	00:00	X	X	X	X	X[d]
C1/D1	EOI-ipi[e]	01:30	X		X		
C2/D1	Predose nivo	00:00					X[f]
C3/D1	Predose nivo	00:00	X	X	X	X	X[f]
C5/D1	Predose nivo	00:00	X	X	X	X	X[f]
C13/D1	Predose nivo	00:00	X	X			
C21/D1	Predose nivo	00:00	X	X			
Every 16 weeks from C21 to 2 years	Predose nivo	00:00	X	X			
Follow-up Samples - Approximately 30 and 100 days from the discontinuation of study drug	NA	NA	X	X	X	X	

[a] If a participant discontinues study drug treatment during the sampling period, they will move to sampling at the follow-up visits.

Abbreviations: C=cycle; D=day; EOI=end of infusion; ipi=ipilimumab; NA=not applicable; nivo=nivolumab

[b] Ipilimumab samples will be collected and only analyzed if needed.

[c] Predose: All predose samples for nivolumab and/or ipilimumab should be taken prior to the start of nivolumab infusion.

[d] Though cabozantinib will be dosed in the evening on Day 1 of Cycle 1, the cabozantinib predose sample on Day 1 can be drawn at the same time nivolumab predose sample is drawn.

[e] EOI-ipi: The end of infusion (EOI) samples for both nivolumab and ipilimumab should be collected immediately (preferably within 2-5 minutes) prior to the end of ipilimumab infusion. If the end of infusion is delayed, the collection of the EOI samples should be delayed accordingly. EOI samples may not be collected from the same IV access as drug was administered, refer to the laboratory manual for additional instructions.

[f] Cabozantinib is preferably dosed at bedtime. A PK sample for cabozantinib should be drawn at the same time when predose PK samples for nivolumab and ipilimumab are drawn as long at the time of the draw for cabozantinib is approximately at least 8 hours after the previous evening dose of cabozanitinib.

9.6 Pharmacodynamics

Refer to Section 9.8.

9.7 Pharmacogenomics

Refer to Section 9.8.6.

9.8 Biomarkers

9.8.1 Additional Research Collection

This protocol will include residual sample storage for additional research (AR).

For All US sites:

Additional research is required for all study participants.

- If the IRB/ethics committees and site agree to the mandatory additional research retention and/or collection, then the study participant must agree to the mandatory additional research as a requirement for inclusion in the study.

For non-US Sites

Additional research is optional for all study participants, except where retention and/or collection is prohibited by local laws or regulations, ethics committees, or institutional requirements.

This collection for additional research is intended to expand the translational R&D capability at Bristol-Myers Squibb, and will support as yet undefined research aims that will advance our understanding of disease and options for treatment. It may also be used to support health authority requests for analysis, and advancement of pharmacodiagnostic development to better target drugs to the right participants. This may also include genetic/genomic exploration aimed at exploring disease pathways, progression and response to treatment etc.

Sample Collection and Storage

All requests for access to samples or data for additional research will be vetted through a diverse committee of the study sponsor's senior leaders in Research and Development (or designee) to ensure the research supports appropriate and well-defined scientific research activities.

- Residual whole blood (DNA), serum, plasma, PBMCs, tumor tissue from whole blood (DNA), serum biomarkers, plasma biomarkers, PBMCs, tumor biopsy collections (see Table 9.8.1-1) will also be retained for additional research purposes

Samples kept for future research will be stored at the BMS Biorepository in Hopewell, NJ, USA or an independent, BMS-approved storage vendor.

- The manager of these samples will ensure they are properly used throughout their usable life and will destroy the samples at the end of the scheduled storage period, no longer than fifteen (15) years after the end of the study or the maximum allowed by applicable law.
- Transfers of samples by research sponsor to third parties will be participant to the recipient's agreement to establish similar storage procedures.

Samples will be stored in a coded fashion, and no researcher will have access to the key. The key is securely held by the Investigator at the clinical site, so there is no direct ability for a researcher to connect a sample to a specific individual.

Further details of sample collection and processing will be provided to the site in the procedure manual.

Table 9.8.1-1: Residual Sample Retention for Additional Research Schedule

Sample Type	Timepoints for which residual samples will be retained
Whole blood (DNA)	All
Serum biomarker	All
Plasma biomarker	All
PBMCs	All
Tumor Biopsy	All

9.8.2 Tissue Specimens

Sufficient tumor tissue specimens, preferably collected within 3 months but no more than 12 months prior to enrollment with an associated pathology report, will be required in the form of a formalin-fixed paraffin-embedded block or 20 unstained slides. A minimum of 10 slides will be acceptable if tumor tissue is limited. In these situations, it is recommended to consult with the protocol team to discuss the specifics of the case. In addition, it is recommended, although optional, for tumor tissue samples to be collected upon progression. These samples may be used for the assessment of markers implicated in resistance.

The tumor tissue samples may be used to assess putative predictive biomarkers of nivolumab, ipilimumab, and cabozantinib efficacy and/or to better characterize the tumor-immune microenvironment. Baseline tumor tissue samples will be evaluated for PD-L1 by IHC during screening by the central lab pathologist, who will determine the number of tumor cells with membranous staining among a minimum of 100 evaluable tumor cells. Participants will be stratified at randomization as either PD-L1 expression $\geq$ 1% versus PD-L1 expression < 1% or indeterminate (ie, sample contains at least 100 evaluable tumor cells but membrane scoring is hampered by high cytoplasmic staining).

Various other markers with potential predictive value for the treatment of RCC with nivolumab, ipilimumab and other immunotherapies are currently under investigation and may be assessed in this study. These tumor tissue biomarkers include, but are not limited to PD-L1, PD-1, PD-L2, TILs or subpopulations of TILs and a Th1 immune mRNA expression signature. Evaluation of MET, AXL, and other markers may be performed by IHC or other methods. Tumor samples may also be used to further characterize the tumor-immune microenvironment, including but not limited to other T cell checkpoint receptors and ligands (eg, Lag-3, Tim-3), and intratumoral immune cell subsets, including T-regulatory cells, myeloid derived suppressor cells, macrophages, natural killer (NK) cells and B cells. These samples may also be used to investigate the effect of nivolumab,

ipilimumab, and cabozantinib on the expression of potentially relevant predictive and/or prognostic RCC biomarkers. Both the baseline tumor sample and the sample collected upon progression may be retrospectively assessed for the expression of other immune related genes, RNAs and/or proteins, or for the presence of immune cell populations using a variety of methodologies inclusive of, but not limited to IHC, RNA Scope, qRT-PCR, RNA seq, genetic mutation detection (including whole exome sequencing) and fluorescent in-situ hybridization (FISH).

9.8.3 Exploratory Serum and Plasma Biomarkers

Blood samples for exploratory serum and plasma biomarker analyses will be drawn at the specified time points. Serum and plasma samples may be assessed by ELISA, seromics, microRNA profiling, circulating tumor DNA measurements, metabolomics and/or other relevant multiplex-based protein assay methods for immune or RCC-related factors that will predict for clinical benefit or correlate with treatment-related adverse events. Potential serum and plasma-based biomarkers to potentially b investigated include, but are not limited to levels of soluble PD-L1, anti-tumor antibodies, cytokines, chemokines, inflammatory factors, angiogenic biomarkers (eg, VEGF-A, soluble VEGF2) NKG2D ligands (eg, soluble MICA), circulating tumor DNA, and microRNAs (such as, but not limited to, miR-513 and miR19b).

9.8.4 Myeloid-Derived Suppressor Cells

Myeloid-derived suppressor cells (MDSCs) are an immune cell population capable of suppressing T cell activation and proliferation. MDSCs will be measured at baseline to assess associations with outcome.

9.8.5 PBMC for Flow Cytometry

Exploratory flow cytometry analysis will be used to assess baseline and on-treatment alterations in the composition/ activation status of cell subsets. Lymphocyte subsets to be assayed may include, but are not limited to, CD4+ and CD8+ subsets (activated, effector/memory, regulatory) and populations of those cells as defined by the expression of activation, exhaustion or signaling markers.

9.8.6 Whole Blood for Genotyping

Whole blood samples for exploratory pharmacogenetic assessment will be collected from all participants. Genomic DNA will be extracted and subsequently assessed for SNPs and other genetic variations in candidate genes that may predispose participants to clinical benefit or adverse events (unless restricted by local requirements). Such genes include, but are not limited to, PD-1, PD-L1, PD-L2 and CTLA-4. Additional use of these data may include correlative analyses aimed at identifying genotypic associations with clinically relevant biomarkers identified by other methodologies described in this section (including whole exome sequencing). Genomic DNA from whole blood will be collected and may be used as a comparator for subjects with tumors examined by whole exome or genome analysis.

9.8.7 Other Assessments

Not applicable.

9.9 Health Economics OR Medical Resource Utilization and Health Economics

Health Economics/Medical Resource Utilization and Health Economics parameters will be evaluated in this study as noted in Section 2.

Medical resource utilization and health economics data, associated with medical encounters, will be collected in the CRF by the investigator and study-site personnel for all participants throughout the study. Protocol-mandated procedures, tests, and encounters are excluded.

The data collected may be used to conduct exploratory economic analyses and will include:

- Number and duration of medical care encounters, including surgeries, and other selected procedures (inpatient and outpatient)
- Duration of hospitalization (total days length of stay, including duration by wards; eg, intensive care unit)
- Number and character of diagnostic and therapeutic tests and procedures
- Outpatient medical encounters and treatments (including physician or emergency room visits, tests and procedures, and medications)].

10 STATISTICAL CONSIDERATIONS

10.1 Sample Size Determination

The sample size of this study accounts for the primary endpoint of progression-free survival (PFS) per BICR in all randomized participants. Assuming a 25% screen failure rate, it is expected that approximately 850 participants will need to be enrolled in order to randomize 638 participants (319 per arm) in a 1:1 ratio. To represent the normal frequency of the favorable risk group in mRCC, the favorable risk participants are capped at approximate 25%; thus, at most 212 favorable risk participants (106 per arm) will be enrolled to randomize 160 favorable risk participants in a 1:1 ratio. The rest of the enrolled participants will provide approximately 478 intermediate/poor risk randomized participants (239 per each arm).

The overall alpha for this study is 0.05 (two-sided). PFS will be evaluated for treatment effect at an alpha of 0.05 (two-sided), with at least 95% power. No interim analysis of PFS is planned. OS will be evaluated for treatment effect at an alpha level of 0.05 (two-sided) with 80% power, accounting for two formal interim analyses to assess efficacy.

Sample Size Justification for Primary PFS Endpoint

The primary endpoint of PFS per BICR of Arm A versus Arm C analysis will be conducted on all randomized participants. The PFS analysis will occur after approximately 9-10 months minimum follow-up on all randomized subjects by which approximately 350 events from Arm A and Arm C are expected. The 350 PFS events provide at least 95% power to detect a HR of 0.68 for PFS of Arm A versus Arm C with a type I error of 0.05 (two-sided). The HR of 0.68 corresponds to a 47% increase in the median PFS, assuming a median PFS of 18.2 months for Arm A and 12.4 months for Arm C. It is projected that an observed HR of 0.811 or less, which corresponds to a 2.89 month or greater improvement in median PFS (12.4 versus 15.3 months), would result in a statistically significant improvement in PFS for the Arm A versus Arm C comparison.

If the formal analysis of PFS among all randomized participants is statistically significant, the formal interim analysis of OS among all randomized participants will be tested, as per hierarchical testing procedure.

Sample Size Computation for Secondary OS Endpoint

The secondary endpoint of OS in all randomized participants specifies the comparison of Arm A versus Arm C. Among all randomized participants, approximately 254 events (ie, deaths) in Arm A and Arm C provides at least 80% power to detect a HR of 0.70 for OS of Arm A and Arm C with an overall type 1 error of 0.05 (two-sided) for each test. The HR of 0.70 corresponds to a 43% increase in the median OS, assuming a median OS of 47.1 months for Arm A and 33 months for Arm C.

Two formal interim analyses of OS are planned for this study. The first interim analysis is planned at the time of final PFS analysis and it is expected to observe 165 OS events (65% of the targeted OS events for final analysis) and the second interim analysis is planned to occur after observing approximately 211 events (83% of targeted OS events needed for final analysis). The stopping boundaries at interim and final analyses will be derived based on the number of deaths using O'Brien and Fleming α spending function. For example, with 165, 211, and 254 observed events in Arm A and Arm C at the first interim, second interim, and final analyses, the respective stopping boundaries would be α=0.011 (two-sided), α=0.025 (two-sided), and α=0.041 (two-sided). If the first interim analysis is performed exactly at 165 deaths, it is projected that an observed HR of 0.673 or less, which corresponds to a 16.0 month or greater improvement in median OS (33 versus 49 months), would result in a statistically significant improvement in OS for the Arm A versus Arm C comparison. At the second interim analysis with 211 deaths, it is projected that an observed HR of 0.734 or less, which corresponds to a 12.0 month or greater improvement in median OS (33 versus 45 months), would result in a statistically significant improvement in OS for the Arm A versus Arm C comparison. At the time of final OS analysis when there are 254 deaths, it is projected that an observed HR of 0.774 or less, which corresponds to a 9.6 month or greater improvement in median OS (33 versus 42.6 months), would result in a statistically significant improvement in OS for the Arm A versus Arm C comparison.

Assuming a constant accrual rate (an average rate of 3 participants/month in the first 4 months, afterwards an average rate of 42 participants/month), the accrual will take approximately 19 months. The final PFS analysis will not occur prior to these conditions being met:

- at least 8 months minimum follow-up on all randomized subjects;
- at least 283 PFS events, which provide at least 90% power to detect a HR of 0.68 for PFS of Arm A versus Arm C; and
- at least 149 OS events, which provide 66% power if the target HR for OS was 0.60. (Note that if the analysis of first interim analysis OS takes place with 149 OS events, the alpha spending for the OS comparison would be 0.007 with a critical HR=0.643.)

This expected PFS analysis will occur at approximately 29 months from FPFV. The second interim and final analyses of OS are expected to occur approximately 34 months and 40 months from FPFV, respectively. Table 10.1-1 summarizes the results of these calculations.

Table 10.1-1: Summary of sample size parameters and schedule of analyses

Primary/Secondary Endpoints	PFS (Primary)	OS (Secondary)[a]
Primary analysis population	All Randomized Participants	
Accrual rate per month for all randomized population	3 participants/month in the first 4 months, afterwards an average rate of 42 participants/month	
Power	95%	80%
Alpha	0.05 2-sided	0.05 2-sided (0.011 at IA1, 0.025 at IA2, 0.041 at FA)
Hypothesized median control vs exp (months)	12.4 vs 18.2	33 vs 47.1
Hypothesized hazard ratio	0.68	0.70
Critical hazard ratio (observed hazard ratio at which a statistically significant difference would be observed) / Difference in median (months) Corresponding to a minimal clinically significant effect size (FA)[b]	0.811 / 2.89	0.774 / 9.6
Critical HR at interim analysis-1(IA1) /effect size	N/A	0.673 / 16.0
Expected number of event for IA1 (percentage of target events)	N/A	165 (65%)
Timing of IA1 from FPFV (months)	N/A	29
Critical HR at interim analysis-2(IA2) /effect size	N/A	0.734 / 12.0
Expected number of event for IA2 (percentage of target events)	N/A	211 (83%)
Timing of IA2 from FPFV (months)	N/A	34
Accrual duration (months)	19	19
Timing of final analysis (FA) from FPFV (months)	29	40
Sample size	638	638

Table 10.1-1: Summary of sample size parameters and schedule of analyses

Primary/Secondary Endpoints	PFS (Primary)	OS (Secondary)[a]
Target number of events (Event Goal)	350	254

[a] OS analyses is participant to significance in hierarchical testing strategy for each pairwise comparison.

[b] The difference in median (months) only applies if the control group median is exactly as hypothesized.

10.2 Populations for Analyses

All analyses will be performed using the treatment arm as randomized (intent to treat), with the exception of dosing and safety, for which the treatment arm as received will be used. For purposes of analysis, the following populations are defined in Table 10.2-1, and all populations for analyses given in this table refer to those participants in Arm A and Arm C. Those participants who randomized to Arm B prior to Revised Protocol 01 will be considered as part of the population of interest for descriptive summary of efficacy and safety analyses.

Table 10.2-1: Populations for Analyses

Population	Description
All Enrolled Participants	All participants who sign informed consent and were registered into the IRT.
All Randomized Participants	All participants who were randomized will be used for the analysis of demography, protocol deviations, baseline characteristics, primary efficacy analysis, secondary efficacy analyses, and outcome research analysis which will be performed for this population.
All Treated Participants	All participants who received at least one dose of any study medication. This is the primary population for exposure and safety analyses.
Pharmacokinetic Participants	All participants with available data from randomized participants dosed with nivolumab, or cabozantinib.
Immunogenicity Participants	All participants with available data from randomized participants dosed with nivolumab.
Biomarker Participants	All participants with available biomarker data from randomized participants.

10.3 Statistical Analyses

The statistical analysis plan will be developed and finalized before database lock and will describe the selection of participants to be included in the analyses, and procedures for accounting for missing, unused, and spurious data. Below is a summary of planned statistical analyses of the primary and secondary endpoints. A description of the participant population will be included in the statistical output reported, including subgroup of age, gender, and race.

10.3.1 Efficacy Analyses

The efficacy analyses will be performed in all randomized participants in Arm A and Arm C. Descriptive summary of efficacy will be provided for participants randomized in Arm B.

Table 10.3.1-1: Efficacy Analyses

Endpoint	Statistical Analysis Methods
Primary	The primary endpoint of this study is to compare PFS per BICR of Arm A (doublet) versus Arm C (single agent), in all randomized participants. PFS is defined as the time between the date of randomization and the first date of the documented progression, or death due to any cause, whichever occurs first. Participants who die without a reported progression (and die without start of subsequent anti-cancer therapy) will be considered to have progressed on the date of their death. Participants who did not progress or die will be censored on the date of their last evaluable tumor assessment on or prior to initiation of subsequent anti-cancer therapy. Participants who did not have any on study tumor assessments and did not die will be censored on their date of randomization. Participants who started anticancer therapy (including tumor-directed radiotherapy and tumor-directed surgery) without a prior reported progression will be censored on the date of their last evaluable tumor assessment on or prior to the initiation of first subsequent anti-cancer therapy. The primary formal comparisons of PFS will be conducted using a two-sided 0.05 stratified log-rank test for each comparison, with IMDC scores, PD-L1 tumor expression, and region at screening per IRT as stratification factors among all randomized participants. Median PFS will be estimated via the Kaplan-Meier product limit method. Two-sided 95% CI for the median PFS will be computed for each randomized arm. Kaplan-Meier plots of PFS will be presented. Hazard ratios (HR) and corresponding two-sided 95% confidence intervals (CI) will be estimated using a Cox proportional hazards model, with treatment arm as a single covariate, stratified by the stratification factors, corresponding to the comparison of PFS. The totality of PFS results will be presented in a single graphical display that includes Kaplan-Meier curves for all treatment arms, the log-rank p-values for the formal comparisons, the HRs and corresponding CIs, and the median PFS estimates and corresponding CIs. The following supportive analyses of PFS will also be conducted: A stratified multivariate Cox regression model will be used in order to estimate the treatment effect after adjustment for possible imbalances in known or potential prognostic factors. The covariates included in this model will be specified in the statistical analysis plan. PFS using an un-stratified log rank test. The hazard ratio associated with treatment will be presented along with the associated two-sided 95% CIs. PFS accounting for missing tumor assessment prior to PFS event (progression or death). This analysis will be performed only if at least 10% of events have missing prior tumor assessment. It will apply the following restriction to the primary definition: If the elapsed time between the PFS event and the last assessment immediately prior to the event is two or more missed visits (more than 12 weeks - 10 days), the participant's PFS will be censored at his/her last tumor assessment prior to the PFS event.
Secondary	The first secondary endpoint is to compare OS of Arm A versus Arm C in all randomized participants. OS is defined as the time between the date of randomization and the date of death due to any cause. A participant who has not died will be censored at the last known alive date. At the time of the primary endpoint analysis, there will be a first interim analysis of OS. The second interim analysis and the final analysis of OS are planned to occur approximately 34 months and 40 months from FPFV, respectively. OS will be compared between these treatment arms using a two sided, 0.05 level log-rank test (adjusted for interim analyses), stratified using IMDC scores, PD-L1 tumor expression, and region at screening as

Table 10.3.1-1: Efficacy Analyses

Endpoint	Statistical Analysis Methods
	stratification factors among all randomized participants. A similar analysis as in PFS will be conducted for OS. Hazard ratio (HR) and corresponding two-sided 95% confidence intervals (CI) will be estimated using a Cox proportional hazards model, with treatment arm as a single covariate, stratified by the stratification factors, corresponding to the comparison of OS. The second secondary endpoint evaluates ORR per BICR for all randomized participants. ORR is defined as the proportion of randomized participants who achieve a best response of complete response (CR) or partial response (PR) using the RECIST 1.1 criteria. Best overall response. (BOR) is defined as the best response designation recorded between the date of randomization and the date of objectively documented progression per RECIST 1.1 or the date of subsequent therapy (including tumor-directed radiotherapy and tumor-directed surgery), whichever occurs first. For participants without document progression or subsequent therapy, all available response designations will contribute to the BOR assessment. Duration of response (DOR) is defined as the time between the date of first confirmed documented response (CR or PR) to the date of first documented tumor progression (per RECIST 1.1) or death due to any cause, whichever occurs first. Participants who neither progress nor die will be censored on the date of their last tumor assessment. Responders who started anti-cancer therapy without a prior reported progression will be censored on the date of their last evaluable tumor assessment prior to the initiation of first subsequent anti-cancer therapy. TTR is defined as the time from randomization to the date of the first confirmed documented response (CR or PR), as assessed by BICR. DOR and TTR will be evaluated for responders (CR or PR) only. Tumor assessments are scheduled to be performed at Week 12 following randomization, every 6 weeks for the first 12 months and then every 12 weeks until progression. ORR will be analyzed at the time of the final PFS analysis. ORRs and corresponding 95% exact CIs will be calculated using the Clopper Pearson method within each treatment arms. An estimate of the response rate and an associated exact two-sided 95% CI (Clopper and Pearson) will be presented, by treatment group. Sensitivity analysis based on investigator-determined ORR may also be performed. DOR and TTR will also be evaluated.
Exploratory	Will be described in the statistical analysis plan finalized before database lock.

10.3.2 Safety Analyses

The safety analyses will be performed in all treated participants in Arm A and Arm C. Participants treated in Arm B will be considered in select safety analyses.

Table 10.3.2-1: Safety Analyses

Endpoint	Statistical Analysis Methods
Primary	The primary endpoint is related to efficacy and there are no primary safety endpoints.
Secondary	Safety will be analyzed at the time of the primary endpoint analysis. Descriptive statistics of safety will be presented using the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) version 4.0 by treatment arm. All AEs, drug-related AEs, SAEs and drug-related SAEs, imAEs, and select AEs will be tabulated using the worst grade per NCI CTCAE version 4.0 criteria by system organ class and preferred term. On-study lab parameters including hematology, coagulation, chemistry, liver function and renal function will be summarized using worse grade per NCI CTCAE version 4.0 criteria.
Exploratory	Will be described in the statistical analysis plan finalized before database lock.

10.3.3 Other Analyses

Pharmacokinetic, pharmacodynamic, and biomarker exploratory analyses will be described in the statistical analysis plan finalized before database lock. The population pharmacokinetics analysis and pharmacodynamic analyses will be presented separately from the main clinical study report.

Immunogenicity may be reported for ADA positive status (such as persistent positive, other positive, only last sample positive, baseline positive) and ADA negative status, relative to baseline. In addition, presence of neutralizing antibody may be reported, if applicable. Effect of immunogenicity on safety/efficacy and biomarkers and PK may be explored.

10.3.3.1 Outcomes Research Analyses

Clinically meaningful score changes (from baseline) reported in the literature were used to define improvement and deterioration. Table 10.3.3.1-1 provides the thresholds for clinically meaningful score changes to be used for the patient-reported outcome scales in this study. For domains without clinically meaningful score changes in the literature, thresholds values were derived using the general guide for calculating thresholds for FACIT instruments provided by Yost and Eton (2005).[56]

Table 10.3.3.1-1: Thresholds Values for Change Scores Judged to be Important to Patients

Instruments and Domains	Thresholds reported in the literature
FKSI-19	
Total Score	--
Disease Related Symptoms Score	2-3[57,58]
Disease Related Symptoms Emotional Score	--
Disease Related Symptoms Physical Score	--
Functional Well Being Score	--
Treatment Side Effects Score	--
EQ-5D-3L[59]	
VAS	7
UK Utility Index	0.08

10.3.4 Interim Analyses

Two interim analyses of OS are planned for this study. The first interim analysis of OS is planned at the time of final PFS analysis and expected after observing 165 deaths (approximately 65% of the targeted OS events) have been observed among all randomized participants in Arm A and Arm C based on above accrual rate and the exponential distribution in each arm. These formal comparisons of OS will allow for early stopping for superiority, and the boundaries for declaring superiority will be derived based on the actual number of deaths using Lan-DeMets spending function with O'Brien and Fleming type of boundary in EAST version 6. If the first interim analysis is performed exactly at 165 deaths, the boundary in terms of statistical significance for declaring superiority would be 0.011 (HR=0.673 with 16 months improvement in median OS for the Arm A versus Arm C comparison (33 versus 49 months)). The second interim analysis of OS

is expected after observing 211 deaths (approximately 83% of the targeted OS events) have been observed among all randomized participants based on above accrual rate and the exponential distribution in each arm. The boundary for declaring superiority in terms of statistical significance for the second interim analysis after 211 events would be 0.025 (HR=0.734 with 12 months improvement in median OS for the doublet versus sunitinib comparison (33 versus 45 months). The boundary for declaring superiority in terms of statistical significance for the final analysis after 254 events would be 0.041 (HR=0.774 with 9.6 months improvement in median OS for the Arm A versus Arm C comparison (33 versus 42.6 months). More details are summarized in Table 10.1-1. An independent statistician external to BMS will perform interim analysis in conjunction with a review by a data monitoring committee.

11 REFERENCES

[1] Znaor, Ariana, et al. International Variations and Trends in Renal Cell Carcinoma Incidence and Mortality. Eur Urol 2015; 67(3): 519-530.

[2] Sun M, Thuret R, Abdollah F, et al. Age-adjusted incidence, mortality, and survival rates of stage-specific renal cell carcinoma in North America: A trend analysis. Eur Urol 2011; 59: 135–41.

[3] Hollingsworth JM, Miller DC, Daignault S, et al. Rising incidence of small renal masses: A need to reassess treatment effect. J Natl Cancer Inst. 2006; 98: 1331–4.

[4] Banumathy G, Cairns P. Signaling Pathways in Renal Cell Carcinoma. Cancer Biol Ther 2010; 10(7): 658–664.

[5] Heng DY, Xie W, Regan MM, et al. External validation and comparison with other models of the International Metastatic Renal-Cell Carcinoma Database Consortium prognostic model: a population-based study. Lancet Oncol 2013; 14(2): 141-8.

[6] NCCN Clinical Practice Guidelines in Oncology; Kidney Cancer. Version 2.2017. Accessed at https://www.nccn.org/professionals/physician_gls/pdf/kidney.pdf on 7-Feb-2017.

[7] Cabometyx (cabozantinib) US Prescribing Information. April 2016.

[8] Liu XD, Hoang A, Zhou L, Kalra S, Yetil A, Sun M, et al. Resistance to Antiangiogenic Therapy Is Associated with an Immunosuppressive Tumor Microenvironment in Metastatic Renal Cell Carcinoma. Canc Immunol Res. 2015; 10: 1017-1029.

[9] Final Study Report CA209016. Bristol-Myers Squibb. Princeton, NJ. December 9, 2016. Document Control No. 930108520.

[10] Atkins MB, Plimack ER, Puzanov I, Fishman MN, McDermott DF, Cho DC, et al. Axitinib in Combination With Pembrolizumab in Patients With Advanced Renal Cell Carcinoma: Preliminary Safety and Efficacy Results [poster no. 773PD]. European Society for Medical Oncology (ESMO), October 7–11, 2016, Copenhagen, Denmark.

[11] Taylor MH, Dutcus CE, Schmidt E, Bagulho T, Li D, Shumaker R, et al. A Phase 1b Trial of Lenvatinib Plus Pembrolizumab in Patients With Selected Solid Tumors [poster no.776PD]. ESMO Annual Conference; October 7–11, 2016; Copenhagen, Denmark.

[12] Apolo AB, Tomita Y, Lee MJ, Lee S, Frosch A, Steinberg SM. Effect of Cabozantinib on immunosuppressive subsets in metastatic urotherlial carcinoma [abstract No. 4501]. 2014. J Clin Oncol. 2014; 32 (5s). 2014 ASCO Annual Meeting.

[13] Pardoll D. Does the immune system see tumors as foreign or self? Annu Rev Immunol 2003; 21: 807-39.

[14] Zitvogel L, Tesniere A, Kroemer G. Cancer despite immunosurveillance: immunoselection and immunosubversion. Nat Rev Immunol 2006; 6: 715-27.

[15] Dunn GP, Bruce AT, Ikeda H, et al. Cancer immunoediting: from immunosurveillance to tumor escape. Nat Immunol 2002; 3: 991-8.

[16] Greenwald RJ, Freeman GH, Sharpe AH. The B7 family revisited. Annu Rev Immunol 2004; 23: 515-48.

[17] Freeman FJ, Long AJ, Iwai Y, et al. Engagement of the PD-1 immunohibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation. J Exp Med 2000; 192(7): 1027-34.

[18] Sharpe AH, Wherry EJ, Ahmed R, et al. The function of programmed cell death 1 and its ligands in regulating autoimmunity and infection. Nature Immunol 2007; 8: 237-45.

[19] Wolchok JD, Hoos A, O'Day S, et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. Clin Cancer Res 2009; 15: 7412-20.

[20] Postow MA, Harding J, Wolchok JD. Targeting immune checkpoints: Releasing the restraints on anti-tumor immunity for participants with melanoma. Cancer J 2012; 18: 153-159.

[21] Investigator Brochure for XL184. Version 12. Exelixis, Inc. South San Francisco, CA. July 2016.

[22] Castellone MD, Carlomagno F, Salvatore G, Santoro M. Receptor tyrosine kinase inhibitors in thyroid cancer. Best Pract Res Clin Endo-crinol Metab 2008; 22: 1023-38

[23] Timar J, Dome B. Antiangiogenic drugs and tyrosine kinases. Anti-cancer Agents Med Chem 2008; 8 :462-9.

[24] McDermott DF, Drake CG, Sznol M, et al. Survival, Durable Response, and Long-Term Safety in Patients With Previously Treated Advanced Renal Cell Carcinoma Receiving Nivolumab. J Clin Oncol. 2015; 33(18): 2013-20.

[25] Yang JC, Hughes M, Kammula U et al. Ipilimumab (Anti-CTLA4 Antibody) causes regression of metastatic renal cell cancer associated with enteritis and hypophysitis. J Immunother 2007; 30(8): 825-830.

[26] Kwilas AR, Ardiani A, Donahue RN, Aftab DT, Hodge JW. Dual effects of a targeted small molecule inhibitor (cabozantinib) on immune mediated killing of tumor cells and immune tumor microenvironment permissiveness when combined with a cancer vaccine. Transl Med 2014; 12: 294.

[27] Tolaney SM, Ziehr DR, Guo H, Ng MR, Barry WT, Higgins MJ, et al. Phase II and Biomarker Study of Cabozantinib in Metastatic Triple Negative Breast Cancer Patients. Oncologist 2016; 21:1-8.

[28] Nivolumab Investigator Brochure. Version 15. Bristol-Myer Squibb Research and Development. Princeton, NJ. June 2016. Document Control No. 930038243.

[29] Ipilimumab Investigator Brochure. Version 19. Bristol-Myer Squibb Research and Development. Princeton, NJ. March 2016. Document Control No. 930017531.

[30] Escudier B, Eisen T, et al. Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2012; 23: 65-71.

[31] Motzer RJ, Hutson TE, Tomczak P, et al. Overall survival and updated results for sunitinib compared with interferon alfa in patients with metastatic renal cell carcinoma. J Clin Oncol 2009; 27: 3584-90.

[32] Motzer RJ, Hutson TE, Cella D, et al. Pazopanib versus sunitinib in metastatic renal-cell carcinoma. N Engl J Med 2013; 369:722-731.

[33] Sutent (sunitinib malate) US Prescribing Information. April 2014.

[34] Motzer RJ, Rini BI, McDermott DF, et al. Nivolumab for Metastatic Renal Cell Carcinoma: Results of a Randomized Phase II Trial. J Clin Oncol 2015; 33(13): 1430-7.

[35] Topalian SL, Hodi FS, Brahmer JR, et al. Safety, Activity, and Immune Correlates of Anti-PD-1 Antibody in Cancer. N Engl J Med. 2012. 366(26): 2443-54.

[36] Motzer RJ, Rini BI, McDermott DF, et al. Nivolumab for Metastatic Renal Cell Carcinoma: Results of a Randomized Phase II Trial. J Clin Oncol 2015; 33(13): 1430-7.

[37] Choueiri TK, Escudier BJ, Kim JJ, et al. LJ. Immunomodulatory activity of nivolumab in previously treated and untreated metastatic renal cell carcinoma (mRCC): Biomarker-based results from a randomized clinical trial. J Clin Oncol 2014; 32:5s (suppl; abstr 5012).

[38] Choueiri TK, Escudier B, Powles T, Mainwaring PN, Rini BI, Donskov F, et al, for the METEOR Investigators. Cabozantinib versus Everolimus in Advanced Renal-Cell Carcinoma. N Engl J Med 2015; 373: 1814-23.

[39] Choueiri TK, Cabozantinib versus everolimus in advanced renal cell carcinoma (METEOR): final results from a randomised, open-label, phase 3 trial. Lancet Oncol 2016; 17(7): 917-27.

[40] Choueiri TK, Halabi S, Sanford BL et al. Cabozantinib Versus Sunitinib As Initial Targeted Therapy for Participants With Metastatic Renal Cell Carcinoma of Poor or Intermediate Risk: The Alliance A031203 CABOSUN Trial J Clin Oncol 2017; 35(6):591-597.

[41] Apolo AB, Mortazavi A, Stein MN, Pal SK, Davarpanah NN, Parnes HL, et al. A Phase I/Expansion Study of Cabozantinib Plus Nivolumab (CaboNivo) Alone or With Ipilimumab (CaboNivoIpi) in Patients With Metastatic Urothelial Carcinoma and other Genitourinary Tumors. ASCO Genitourinary Cancers Symposium, Orlando, Florida. February, 2017.

[42] Heng DYC, Xie W, Regan MM, Harshman LC, Bjarnason GA, Vaishampayan UN, et al. External validation and comparison with other models of the International Metastatic Renal-Cell Carcinoma Database Consortium prognostic model: a population-based study. Lancet Oncol 2013; 14: 141-48.

[43] Escudier B, Tannir NM, McDermott DF et al. CheckMate 214: Efficacy and Safety of Nivolumab Plus Ipilimumab vs Sunitinib for Treatment-Naïve Advanced or Metastatic Renal Cell Carcinoma, Including IMDC Risk and PD-L1 Expression Subgroups, Oral presentation presented at: ESMO Congress Madrid, Spain 2017.

[44] Schadendorf D et al Efficacy and quality of life outcomes in patients with advanced melanoma (MEL) who discontinued treatment with nivolumab (NIVO) plus ipilimumab (IPI) due to toxicity in a phase III trial (CheckMate 067). 2016 Annual Meeting of the European Association of Dermato-Oncology; August 31-September 3, 2016; Vienna, Austria.

[45] Schadendorf D, et al. Pooled Analysis of Long-Term Survival Data from Phase II and Phase III Trials of Ipilimumab in Unresectable or Metastatic Melanoma. J Clin Oncol 2015; 33: 1889-94.

[46] Brahmer J, et al. Five-year follow-up from the CA209-003 study of nivolumab in previously treated advanced non-small cell lung cancer: clinical characteristics of long-term survivors. Oral presentation presented at: American Association for Cancer Research (AACR) Annual Meeting; April 1-5, 2017; Washington, DC, USA.

[47] Felip E et al., Three-year follow-up from Checkmate 017/057: Nivolumab versus docetaxel in patients with previously treated advanced non-small lung cancer (NSCLC). Poster discussion presentation at the European Society of Medical Oncology Annual Meeting. 2017 Sep 8-12; Madrid, Spain. Poster 1301PD.

[48] Herbst RS et al. KEYNOTE-010: Durable clinical benefit in patients with previously treated, PD-L1-expressing NSCLC who completed pembrolizumab. Poster presentation at the World Conference on Lung Cancer 2016 Dec 4-7; Vienna, Austria.

[49] Robert, C et al. Long-term outcomes in patients with ipilimumab-naive advanced melanoma in the phase 3 KEYNOTE-006 study who completed pembrolizumab treatment; Journal of Clinical Oncology 2017 35:15_suppl, 9504-9504.

[50] Spigel DR, et al. Randomized results of continuous vs 1-year fixed-duration in patients with advanced non-small cell lung cancer. Oral presentation at the European Society of Medical Oncology Annual Meeting. 2017 Sep 8-12; Madrid, Spain.

[51] Bennouna J, Sastre J, Arnold D, et al. Continuation of bevacizumab after first progression in metastatic colorectal cancer (ML18147): a randomised phase 3 trial. Lancet Oncol. 2013; 14:29-37.

[52] Takeda M, Yamanaka T, Seto T, et al. Bevacizumab beyond disease progression after first-line treatment with bevacizumab plus chemotherapy in advanced nonsquamous non–small cell lung cancer (West Japan Oncology Group 5910L): an open-label, randomized, phase 2 trial. Cancer 2016; 122:1050-1059.

[53] Rao D, Butt Z, Rosenbloom S, et al. A comparison of the Renal Cell Carcinoma-Symptom Index (RCC-SI) and the Functional Assessment of Cancer Therapy-Kidney Symptom Index (FKSI). J Pain Symptom Manage 2009;38:291-98.

[54] Rothrock NE, Jensen SE, Beaumont JL, et al. Development and initial validation of the NCCN/FACT symptom index for advanced kidney cancer. Value Health 2013;16:789-96.

[55] EuroQol Group. EuroQol--A new facility for the measurement of health-related quality of life. Health Policy 1990;16:199–208.

[56] Yost KJ and Eton DT. Combining distribution and anchor-based approaches to determine minimally important differences: the FACIT experience. Eval Health Prof 2005;28(2):172-191.

[57] Motzer RJ, Escudier B, Oudard S et al. Efficacy of everolimus in advanced renal cell carcinoma: a double-blind, randomised, placebo-controlled phase III trial. Lancet 2008;372(9637):449–456.

[58] Cella D, Yount S, Brucker PS et al. Development and validation of a scale to measure disease-related symptoms of kidney cancer. Value Health 2007;10:285–293.

[59] Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. Health Qual Life Outcomes 2007;5:70.

12 APPENDICES

APPENDIX 1 ABBREVIATIONS AND TRADEMARKS

Term	Definition
AE	adverse event
ACLS	advanced cardiac life support
AHA	alpha hydroxy acid
AI	accumulation index
AIDS	acquired immunodeficiency syndrome
AJCC	American Joint Committee on Cancer
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANC	absolute neutrophil count
ANOVA	analysis of variance
APC	antigen-presenting cell
aPTT	activated partial thromboplastin time
AST	aspartate aminotransferase
AT	aminotransaminases
AUC	area under the concentration-time curve
AUC(INF)	area under the concentration-time curve from time zero extrapolated to infinite time
AUC(0-T)	area under the concentration-time curve from time zero to the time of the last quantifiable concentration
AUC(TAU)	area under the concentration-time curve in one dosing interval
A-V	atrioventricular
AXL	member of the TAM (Tyro-3, Axl, Mer) receptor tyrosine kinases (RTK) subfamily
β-HCG	beta-human chorionic gonadotrophin
Bcl-xL	anti-apoptotic member of the B-cell lymphoma 2 (BCL-2) protein family
BICR	blinded independent central review
BID, bid	bis in die, twice daily
BLQ	below limit of quantification
BMI	body mass index
BMS	Bristol-Myers Squibb

Term	Definition
BOR	best overall response
BP	blood pressure
BTLA	B and T lymphocyte attenuator
BUN	blood urea nitrogen
C	Celsius
C	cycle
C12	concentration at 12 hours
C24	concentration at 24 hours
Ca, Ca^{++}	calcium
Cavg	average concentration
Cavgss	average concentration at steady state
CBC	complete blood count
CD	cluster of differentiation
CD3, CD8, CD14, CD28	cluster of differentiation 3, 8, 14, 28
CD137	member of the tumor necrosis factor (TNF) receptor family. Alternative names are TNF receptor superfamily member 9 (TNFRSF9), 4-1BB and induced by lymphocyte activation (ILA)
Cexpected-tau	expected concentration in a dosing interval
CFR	Code of Federal Regulations
CHF	congestive heart failure
CI	confidence interval
Cl, $C1^{-}$	chloride
CLcr	creatinine clearance
CLT	total body clearance
CLT/F (or CLT)	apparent total body clearance
CLT/F/fu or CLT/fu	Apparent clearance of free drug or clearance of free if (if IV)
cm	centimeter
Cmax, CMAX	maximum observed concentration

Term	Definition
c-MET	tyrosine-protein kinase mesenchymal-epithelial transition (MET) or hepatocyte growth factor receptor, is a protein that in humans is encoded by the MET gene
CMH	Cochran-Mantel-Haenzel
Cmin, CMIN	minimum observed concentration
CMV	cytomegalovirus
CNS	central nervous system
CONSORT	Consolidated standards of reporting trials
CR	complete response
CRC	Clinical Research Center
CRF	Case Report Form, paper or electronic
CT	computed tomography
CTAg	clinical trial agreement
CTCAE	Common Terminology Criteria for Adverse Events
CTLA-4	cytotoxic T lymphocyte antigen-4
Ctrough	Trough observed plasma concentration
CV	coefficient of variation
CVA	cerebrovascular accident
CYP	cytochrome p-450
D	day
DBP	diastolic blood pressure
D/C	discontinue
DILI	drug induced liver injury
dL	deciliter
DMC	data monitoring committee
DNA	deoxyribonucleic acid
DOR	duration of response
DSM IV	Diagnostic and Statistical Manual of Mental Disorders (4^{th} Edition)
DVT	deep vein thrombosis
EA	extent of absorption
EC50	half maximal effective concentration

Term	Definition
ECG	electrocardiogram
eCRF	Electronic Case Report Form
EDC	Electronic Data Capture
EEG	electroencephalogram
e.g., eg	exempli gratia (for example)
eGFR	estimated glomerular filtration rate
ELISA	enzyme-linked immunosorbent assay
EOI	end of infusion
EQ-5D-3L	EuroQoL Group's instrument to measure general health status
ESR	Expedited Safety Report
F	bioavailability
FA	final analysis
FDA	Food and Drug Administration
FDG-PET	fludeoxyglucose- positron emission tomography
FFPE	formalin-fixed, paraffin-embedded
FISH	fluorescent in-situ hybridization
FKSI-19	Functional Assessment of Cancer Therapy - Kidney Symptom Index
FLT-3	Fms-related tyrosine kinase 3
FSH	follicle stimulating hormone
fT3, fT4	free thyroxine (fT3), free triiodothyronine (fT4)
FU	follow up
g	gram
GC	gas chromatography
GCP	Good Clinical Practice
GGT	gamma-glutamyl transferase
GFR	glomerular filtration rate
GI	gastrointestinal
h	hour
HA	health authorities
HBsAg	hepatitis B surface antigen

Term	Definition
HBV	hepatitis B virus
hCG, HCG	human chorionic gonadotropin
HCO_3^-	bicarbonate
HCV	hepatitis C virus
HFS	hand foot syndrome
HIV	Human Immunodeficiency Virus
HR	heart rate, hazard ratio
HrQoL	health-related quality of life
HRT	hormone replacement therapy
ICD	International Classification of Diseases
IC50	half maximal inhibitory concentration
ICH	International Conference on Harmonisation
ICOS	inducible co-stimulator
i.e., ie	id est (that is)
IEC	Independent Ethics Committee
IFN-γ	interferon-γ
IFN-alpha	interferon alphas
IgG1	immunoglobulin G1
IHC	Immunohistochemistry
IL	interleukin
IL-2	interleukin-2
IMAEs, imAEs	immune-mediated adverse events
IMDC	International Metastatic Renal Cell Carcinoma Database Consortium
IMP	investigational medicinal products
IND	Investigational New Drug Exemption
INR	international normalized ratio
IP	investigational product
ipi	ipilimumab
iRAEs	immune-related adverse events
IRB	Institutional Review Board

Term	Definition
IRC	independent radiologic review committee
IRT	Interactive Response Technology
IU	International Unit
IV	intravenous
K	slope of the terminal phase of the log concentration-time curve
K_3EDTA	potassium ethylenediaminetetraacetic acid
K, K^+	potassium
kg	kilogram
KIT	platelet-derived growth factor receptors (PDGFRs)
KPS	Karnofsky Performance Status
L	liter
LAM	Lactation amenorrhea method
LC	liquid chromatography
LC-MS-MS	liquid chromatography tandem-mass spectrometry
LDH	lactate dehydrogenase
LFT	liver function test
LINAC	linear accelerator
LLN	lower limit of normal
LMWH	low molecular weight heparin
ln	natural logarithm
MDSCs	myeloid-derived suppressor cells
MER	proto-oncogene tyrosine-protein kinase MER is an enzyme that in humans is encoded by the MERTK gene
MET	mesenchymal-epithelial transition factor, a tyrosine kinase receptor
mg	milligram
Mg, Mg^{++}	magnesium
MHC	major histocompatibility complex
min	minute
mL	milliliter
MLR	mixed lymphocyte reaction

Term	Definition
mmHg	millimeters of mercury
mOS	median overall survival
mRCC	metastatic renal cell carcinoma
MRI	magnetic resonance imaging
MS	mass spectrometry
MSKCC	Memorial Sloan Kettering Cancer Center
MTD	maximum tolerated dose
mTOR	mammalian target of rapamycin
mUC	metastatic urothelial carcinoma
MUGA	multigated acquisition scan
mWHO	modified World Health Organization
μg	microgram
N	number of subjects or observations
N/A, NA	not applicable
Na, Na^+	sodium
NaF	Sodium fluoride
NCCN	National Comprehensive Cancer Network
NCI	National Cancer Institute
NE	not evaulable
ng	nanogram
NIMP	non-investigational medicinal products
nivo	nivolumab
NK	natural killer
NKG2D	encoded by KLRK1 gene which is located in the NK-gene complex (NKC)
NR	not reached
NSAID	nonsteroidal anti-inflammatory drug
NSCLC	non-small cell lung cancer
ORR	objective response rate
OS	overall survival

Term	Definition
P	phosporous
PCR	polyermase chain reaction
PD	progressive disease, disease progression
PD	pharmacodynamics
PD-1	programmed death-1
PD-L1	programmed death-ligand 1
PE	pulmonary embolism
PET	positron emission tomography
PFS	progression-free survival
PK	pharmacokinetics
PO	per os (by mouth route of administration)
PR	partial response
PT	prothrombin time
PTT	partial thromboplastin time
Q2W, Q3W	every 2 weeks, every 3 weeks
QC	quality control
QD, qd	quaque die, once daily
qRT-PCR	quantitative real time polyermase chain reaction
QTcF	Fridericia corrected QT
R^2	coefficient of determination
RBC	red blood cell
RCC	renal cell carcinoma
RECIST	Response Evaluation Criteria in Solid Tumors
RET	proto-oncogene encodes a receptor tyrosine kinase for members of the glial cell line-derived neurotrophic factor family of extracellular signalling molecules
RNA	ribonucleic acid
ROS1	proto-oncogene tyrosine-protein kinase ROS is an enzyme that in humans is encoded by the ROS1 gene
ROW	rest of the world
RR	respiratory rate

Term	Definition
RS	radiosurgery
RTK	receptor tyrosine kinases
SAE	serious adverse event
SBP	systolic blood pressure
SCC	squamous cell carcinoma
SD	standard deviation, stable disease
SNP	single-nucleotide polymorphism
SOP	Standard Operating Procedures
Subj	subject
SVC	superior vena cava
t	temperature
T	time
TAMs	tumor-assisted macrophages
TAO	Trial Access Online, the BMS implementation of an EDC capability
TCR	T-cell receptor
T-HALF	Half life
TID, tid	ter in die, three times a day
TIE-2	tunica interna endothelial cell kinase 2
TIL	tumour-infiltrating lymphocyte,
TKIs	tyrosine kinase inhibitors
Tmax, TMAX	time of maximum observed concentration
Tregs	regulatory T cells
TRKB	tropomyosin receptor kinase B (TrkB), also known as tyrosine receptor kinase B
TSH	thyroid stimulating hormone
TTR	time to response
TYRO3	tyrosine-protein kinase receptor TYRO3 is an enzyme that in humans is encoded by the TYRO3 gene
ULN	upper limit of normal
UPCR	urine protein to creatinine ratio
UV	ultraviolet

Term	Definition
VAS	visual analog rating scale
VEGF	vascular endothelial growth factor
VEGFR2	vascular endothelial growth factor receptor 2
Vss/F (or Vss)	apparent volume of distribution at steady state
Vz	volume of distribution of terminal phase (if IV and if multi-exponential decline)
W	washout
WBC	white blood cell
WHO	World Health Organization
wks	weeks
WOCBP	women of childbearing potential
WNOCBP	women **not** of childbearing potential
x g	times gravity
XL184	cabozantinib

APPENDIX 2 STUDY GOVERANCE CONSIDERATIONS

The term 'Participant' is used in the protocol to refer to a person who has consented to participate in the clinical research study. The term 'Subject' used in the CRF is intended to refer to a person (Participant) who has consented to participate in the clinical research study.

REGULATORY AND ETHICAL CONSIDERATIONS

GOOD CLINICAL PRACTICE

This study will be conducted in accordance with:

- Consensus ethical principles derived from international guidelines including the Declaration of Helsinki and Council for International Organizations of Medical Sciences (CIOMS) International Ethical Guidelines Good Clinical Practice (GCP),
- as defined by the International Council on Harmonisation (ICH)
- in accordance with the ethical principles underlying European Union Directive 2001/20/EC
- United States Code of Federal Regulations, Title 21, Part 50 (21CFR50)
- applicable local requirements.

The study will be conducted in compliance with the protocol. The protocol and any amendments and the participant informed consent will receive approval/favorable opinion by Institutional Review Board/Independent Ethics Committee (IRB/IEC), and regulatory authorities according to applicable local regulations prior to initiation of the study.

All potential serious breaches must be reported to the Sponsor or designee immediately. A potential serious breach is defined as a Quality Issue (eg, protocol deviation, etc) that is likely to affect, to a significant degree one or more of the following: (1) the physical, safety or mental integrity of one or more subjects/participants; (2) the scientific value of the trial (eg, reliability and robustness of generated data). Items (1) or (2) can be associated with either GCP Regulation(s) or Trial protocol(s).

Personnel involved in conducting this study will be qualified by education, training, and experience to perform their respective tasks.

This study will not use the services of study personnel where sanctions have been invoked or where there has been scientific misconduct or fraud (eg, loss of medical licensure, debarment).

INSTITUTIONAL REVIEW BOARD/INDEPENDENT ETHICS COMMITTEE

Before study initiation, the investigator must have written and dated approval/favorable opinion from the IRB/IEC for the protocol, consent form, participant recruitment materials (eg, advertisements), and any other written information to be provided to subjects/participants. The investigator or BMS should also provide the IRB/IEC with a copy of the Investigator Brochure or product labeling information to be provided to subjects/participants and any updates.

The investigator, Sponsor or designee should provide the IRB/IEC with reports, updates and other information (eg, expedited safety reports, amendments, and administrative letters) according to regulatory requirements or institution procedures.

COMPLIANCE WITH THE PROTOCOL AND PROTOCOL REVISIONS

The investigator should not implement any deviation or change to the protocol without prior review and documented approval/favorable opinion of an amendment from the IRB/IEC (and if applicable, also by local health authority) except where necessary to eliminate an immediate hazard(s) to study subjects/participants.

If a deviation or change to a protocol is implemented to eliminate an immediate hazard(s) prior to obtaining relevant approval/favorable opinion(s) the deviation or change will be submitted, as soon as possible to:

- IRB/IEC
- Regulatory Authority(ies), if applicable by local regulations (per national requirements)

Documentation of approval/favorable opinion signed by the chairperson or designee of the IRB(s)/IEC(s) and if applicable, also by local health authority must be sent to BMS.

If an amendment substantially alters the study design or increases the potential risk to the participant: (1) the consent form must be revised and submitted to the IRB(s)/IEC(s) for review and approval/favorable opinion; (2) the revised form must be used to obtain consent from subjects//participants currently enrolled in the study if they are affected by the amendment; and (3) the new form must be used to obtain consent from new subjects/participants prior to enrollment.

If the revision is done via an administrative letter, investigators must inform their IRB(s)/IEC(s).

FINANCIAL DISCLOSURE

Investigators and sub-Investigators will provide the Sponsor with sufficient, accurate financial information in accordance with local regulations to allow the Sponsor to submit complete and accurate financial certification or disclosure statements to the appropriate health authorities. Investigators are responsible for providing information on financial interests during the course of the study and for 1 year after completion of the study.

INFORMED CONSENT PROCESS

Investigators must ensure that subjects/participants are clearly and fully informed about the purpose, potential risks, and other critical issues regarding clinical studies in which they volunteer to participate.

In situations where consent cannot be given to subjects/participants, their legally acceptable representatives (as per country guidelines) are clearly and fully informed about the purpose, potential risks, and other critical issues regarding clinical studies in which the participant volunteers to participate.

Sponsor or designee will provide the investigator with an appropriate (ie, Global or Local) sample informed consent form which will include all elements required by ICH, GCP and applicable regulatory requirements. The sample informed consent form will adhere to the ethical principles that have their origin in the Declaration of Helsinki.

Investigators must:

- Provide a copy of the consent form and written information about the study in the language in which the participant is most proficient prior to clinical study participation. The language must be non-technical and easily understood.
- Allow time necessary for participant or participant's legally acceptable representative to inquire about the details of the study.
- Obtain an informed consent signed and personally dated by the participant or the participant's legally acceptable representative and by the person who conducted the informed consent discussion.
- Obtain the IRB/IEC's written approval/favorable opinion of the written informed consent form and any other information to be provided to the subjects/participants, prior to the beginning of the study, and after any revisions are completed for new information.

If informed consent is initially given by a participant's legally acceptable representative or legal guardian, and the participant subsequently becomes capable of making and communicating his or her informed consent during the study, consent must additionally be obtained from the participant.

Revise the informed consent whenever important new information becomes available that is relevant to the participant's consent. The investigator, or a person designated by the investigator, should fully inform the participant or the participant's legally acceptable representative or legal guardian, of all pertinent aspects of the study and of any new information relevant to the participant's willingness to continue participation in the study. This communication should be documented.

The confidentiality of records that could identify subjects/participants must be protected, respecting the privacy and confidentiality rules applicable to regulatory requirements, the subjects'/participants' signed ICF and, in the US, the subjects'/participants' signed HIPAA Authorization.

The consent form must also include a statement that BMS and regulatory authorities have direct access to participant records.

Subjects/participants unable to give their written consent (eg, stroke or subjects/participants with or severe dementia) may only be enrolled in the study with the consent of a legally acceptable representative. The participant must also be informed about the nature of the study to the extent compatible with his or her understanding, and should this participant become capable, he or she should personally sign and date the consent form as soon as possible. The explicit wish of a participant who is unable to give his or her written consent, but who is capable of forming an

opinion and assessing information to refuse participation in, or to be withdrawn from, the clinical study at any time should be considered by the investigator.

The rights, safety, and well-being of the study subjects/participants are the most important considerations and should prevail over interests of science and society.

SOURCE DOCUMENTS

The Investigator is responsible for ensuring that the source data are accurate, legible, contemporaneous, original and attributable, whether the data are hand-written on paper or entered electronically. If source data are created (first entered), modified, maintained, archived, retrieved, or transmitted electronically via computerized systems (and/or any other kind of electronic devices) as part of regulated clinical trial activities, such systems must be compliant with all applicable laws and regulations governing use of electronic records and/or electronic signatures. Such systems may include, but are not limited to, electronic medical/health records (EMRs/EHRs), adverse event tracking/reporting, protocol required assessments, and/or drug accountability records).

When paper records from such systems are used in place of electronic format to perform regulated activities, such paper records should be certified copies. A certified copy consists of a copy of original information that has been verified, as indicated by a dated signature, as an exact copy having all of the same attributes and information as the original.

STUDY TREATMENT RECORDS

Records for study treatments (whether supplied by BMS, its vendors, or the site) must substantiate study treatment integrity and traceability from receipt, preparation, administration, and through destruction or return. Records must be made available for review at the request of BMS/designee or a Health Authority.

If	Then
Supplied by BMS (or its vendors):	Records or logs must comply with applicable regulations and guidelines and should include: • amount received and placed in storage area • amount currently in storage area • label identification number or batch number • amount dispensed to and returned by each participant, including unique participant identifiers • amount transferred to another area/site for dispensing or storage • nonstudy disposition (eg, lost, wasted) • amount destroyed at study site, if applicable • amount returned to BMS • retain samples for bioavailability/bioequivalence/biocomparability, if applicable • dates and initials of person responsible for Investigational Product dispensing/accountability, as per the Delegation of Authority Form.
Sourced by site, and not supplied by BMS or its vendors (examples include IP sourced from the sites stock or commercial supply, or a specialty pharmacy)	The investigator or designee accepts responsibility for documenting traceability and study drug integrity in accordance with requirements applicable under law and the SOPs/standards of the sourcing pharmacy.

BMS or designee will provide forms to facilitate inventory control if the investigational site does not have an established system that meets these requirements.

CASE REPORT FORMS

An investigator is required to prepare and maintain adequate and accurate case histories designed to record all observations and other data pertinent to the investigation on each individual treated or entered as a control in the investigation. Data that are derived from source documents and reported on the CRF must be consistent with the source documents or the discrepancies must be explained. Additional clinical information may be collected and analyzed in an effort to enhance understanding of product safety. CRFs may be requested for AEs and/or laboratory abnormalities that are reported or identified during the course of the study.

For sites using the Sponsor or designee electronic data capture tool, electronic CRFs will be prepared for all data collection fields except for fields specific to SAEs and pregnancy, which will be reported on the electronic SAE form and Pregnancy Surveillance form, respectively. If electronic SAE form is not available, a paper SAE form can be used. Spaces may be left blank

only in those circumstances permitted by study-specific CRF completion guidelines provided by Sponsor or designee.

The confidentiality of records that could identify subjects/participants must be protected, respecting the privacy and confidentiality rules in accordance with the applicable regulatory requirement(s).

The investigator will maintain a signature sheet to document signatures and initials of all persons authorized to make entries and/or corrections on CRFs.

The completed CRF, SAE/pregnancy CRFs, must be promptly reviewed, signed, and dated by the investigator or qualified physician who is a subinvestigator and who is delegated this task on the Delegation of Authority Form. Subinvestigators in Japan may not be delegated the CRF approval task For electronic CRFs, review and approval/signature is completed electronically through the BMS electronic data capture tool. The investigator must retain a copy of the CRFs including records of the changes and corrections.

Each individual electronically signing electronic CRFs must meet Sponsor or designee training requirements and must only access the BMS electronic data capture tool using the unique user account provided by Sponsor or designee. User accounts are not to be shared or reassigned to other individuals

MONITORING

Sponsor or designee representatives will review data centrally to identify potential issues to determine a schedule of on-site visits for targeted review of study records.

Representatives of BMS must be allowed to visit all study site locations periodically to assess the data quality and study integrity. On site they will review study records and directly compare them with source documents, discuss the conduct of the study with the investigator, and verify that the facilities remain acceptable. Certain CRF pages and/or electronic files may serve as the source documents:

In addition, the study may be evaluated by Sponsor or designee internal auditors and government inspectors who must be allowed access to CRFs, source documents, other study files, and study facilities. BMS audit reports will be kept confidential.

The investigator must notify BMS promptly of any inspections scheduled by regulatory authorities, and promptly forward copies of inspection reports to Sponsor or designee.

RECORDS RETENTION

The investigator (or head of the study site in Japan) must retain all study records and source documents for the maximum period required by applicable regulations and guidelines, or institution procedures, or for the period specified by BMS or designee, whichever is longer. The investigator (or head of the study site in Japan) must contact BMS prior to destroying any records associated with the study.

BMS or designee will notify the investigator (or head of the study site in Japan) when the study records are no longer needed.

If the investigator withdraws from the study (eg, relocation, retirement), the records shall be transferred to a mutually agreed upon designee (eg, another investigator, study site, IRB). Notice of such transfer will be given in writing to BMS or designee.

RETURN OF STUDY TREATMENT

For this study, study treatments (those supplied by BMS, a vendor or sourced by the investigator) such as partially used study treatment containers, vials and syringes may be destroyed on site.

If..	Then
Study treatments supplied by BMS (including its vendors	Any unused study treatments supplied by BMS can only be destroyed after being inspected and reconciled by the responsible Study Monitor unless study treatments containers must be immediately destroyed as required for safety, or to meet local regulations (eg, cytotoxics or biologics). If study treatments will be returned, the return will be arranged by the responsible Study Monitor.
Study treatments sourced by site, not supplied by BMS (or its vendors) (examples include study treatments sourced from the sites stock or commercial supply, or a specialty pharmacy)	It is the investigator's or designee's responsibility to dispose of all containers according to the institutional guidelines and procedures.

It is the investigator's or designee's responsibility to arrange for disposal, provided that procedures for proper disposal have been established according to applicable federal, state, local, and institutional guidelines and procedures, and provided that appropriate records of disposal are kept. The following minimal standards must be met:

- On-site disposal practices must not expose humans to risks from the drug.
- On-site disposal practices and procedures are in agreement with applicable laws and regulations, including any special requirements for controlled or hazardous substances.
- Written procedures for on-site disposal are available and followed. The procedures must be filed with the site's SOPs and a copy provided to BMS upon request.
- Records are maintained that allow for traceability of each container, including the date disposed of, quantity disposed, and identification of the person disposing the containers. The method of disposal, ie, incinerator, licensed sanitary landfill, or licensed waste disposal vendor must be documented.
- Accountability and disposal records are complete, up-to-date, and available for the Monitor to review throughout the clinical trial period.

It is the investigator's or designee's responsibility to arrange for disposal of all empty containers.

If conditions for destruction cannot be met the responsible Study Monitor will make arrangements for return of study treatments provided by BMS (or its vendors). Destruction of non- study treatments sourced by the site, not supplied by BMS, is solely the responsibility of the investigator or designee.

CLINICAL STUDY REPORT

A Signatory Investigator must be selected to sign the clinical study report.

For each CSR related to this protocol, the following criteria will be used to select the signatory investigator:

- National Coordinating Investigator
- Study Steering Committee chair or their designee
- Participant recruitment (eg, among the top quartile of enrollers)
- Involvement in trial design
- Regional representation (eg, among top quartile of enrollers from a specified region or country)
- Other criteria (as determined by the study team)

SCIENTIFIC PUBLICATIONS

The data collected during this study are confidential and proprietary to Sponsor or designee. Any publications or abstracts arising from this study must adhere to the publication requirements set forth in the clinical trial agreement (CTAg) governing [Study site or Investigator] participation in the study. These requirements include, but are not limited to, submitting proposed publications to Sponsor or designee at the earliest practicable time prior to submission or presentation and otherwise within the time period set forth in the CTAg.

Scientific Publications (such as abstracts, congress podium presentations and posters, and manuscripts) of the study results will be a collaborative effort between the study Sponsor and the external authors. No public presentation or publication of any interim results may be made by any principal investigator, sub-investigator or any other member of the study staff without the prior written consent of the Sponsor.

Authorship of publications at BMS is aligned with the criteria of the International Committee of Medical Journal Editors (ICMJE, www.icmje.org). Authorship selection is based upon significant contributions to the study (ie, ICMJE criterion #1). Authors must meet all 4 ICMJE criteria for authorship:

1) Substantial intellectual contribution to the conception or design of the work; or the acquisition of data (ie, evaluable subjects with quality data), analysis, or interpretation of data for the work (eg, problem solving, advice, evaluation, insights and conclusion); AND
2) Drafting the work or revising it critically for important intellectual content; AND
3) Final approval of the version to be published; AND

4) Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Those who make the most significant contributions, as defined above, will be considered by BMS for authorship of the primary publication. Sub-investigators will generally not be considered for authorship in the primary publication. Geographic representation will also be considered.

Authors will be listed by order of significant contributions (highest to lowest), with the exception of the last author. Authors in first and last position have provided the most significant contributions to the work.

For secondary analyses and related publications, author list and author order may vary from primary to reflect additional contributions.

APPENDIX 3 ADVERSE EVENTS AND SERIOUS ADVERSE EVENTS: DEFINITIONS AND PROCEDURES FOR RECORDING, EVALUATING, FOLLOW UP AND REPORTING

ADVERSE EVENTS

Adverse Event Definition:
An Adverse Event (AE) is defined as any new untoward medical occurrence or worsening of a preexisting medical condition in a clinical investigation participant administered study drug and that does not necessarily have a causal relationship with this treatment.
An AE can therefore be any unfavorable and unintended sign (such as an abnormal laboratory finding), symptom, or disease temporally associated with the use of study drug, whether or not considered related to the study drug.

SERIOUS ADVERSE EVENTS

Serious Adverse Event (SAE) is defined as any untoward medical occurrence that, at any dose:
Results in death
Is life-threatening (defined as an event in which the participant was at risk of death at the time of the event; it does not refer to an event which hypothetically might have caused death if it were more severe)
Requires inpatient hospitalization or causes prolongation of existing hospitalization (see NOTE below) NOTE: The following hospitalizations are not considered SAEs in BMS clinical studies: o a visit to the emergency room or other hospital department < 24 hours, that does not result in admission (unless considered an important medical or life-threatening event) o elective surgery, planned prior to signing consent o admissions as per protocol for a planned medical/surgical procedure o routine health assessment requiring admission for baseline/trending of health status (e.g., routine colonoscopy) o medical/surgical admission other than to remedy ill health and planned prior to entry into the study. Appropriate documentation is required in these cases o admission encountered for another life circumstance that carries no bearing on health status and requires no medical/surgical intervention (e.g., lack of housing, economic inadequacy, caregiver respite, family circumstances, administrative reason) o admission for administration of anticancer therapy in the absence of any other SAEs (applies to oncology protocols)
Results in persistent or significant disability/incapacity
Is a congenital anomaly/birth defect
is an important medical event (defined as a medical event(s) that may not be immediately life-threatening or result in death or hospitalization but, based upon appropriate medical and scientific judgment, may jeopardize the participant or may require intervention [e.g., medical, surgical] to prevent one of the other serious outcomes listed in the definition above.) Examples of such events include, but are not limited to, intensive treatment in an emergency room or at home for allergic bronchospasm; blood dyscrasias or convulsions that do not result in hospitalization.) Potential drug induced liver injury (DILI) is also considered an important medical event. (See Section 9.2.7 for the definition of potential DILI.)

Suspected transmission of an infectious agent (e.g., pathogenic or nonpathogenic) via the study treatment is an SAE.

Although pregnancy, overdose, cancer, and potential drug induced liver injury (DILI) are not always serious by regulatory definition, these events must be handled as SAEs. (See Section 9.2.5 for reporting pregnancies).

Any component of a study endpoint that is considered related to study therapy should be reported as SAE (e.g., death is an endpoint, if death occurred due to anaphylaxis, anaphylaxis must be reported).

EVALUATING AES AND SAES

Assessment of Causality
The causal relationship to study drug is determined by a physician and should be used to assess all adverse events (AE). The causal relationship can be one of the following: Related: There is a reasonable causal relationship between study drug administration and the AE. Not related: There is not a reasonable causal relationship between study drug administration and the AE. The term "reasonable causal relationship" means there is evidence to suggest a causal relationship.

Follow-up of AEs and SAEs
If only limited information is initially available, follow-up reports are required. (Note: Follow-up SAE reports must include the same investigator term(s) initially reported.) If an ongoing SAE changes in its intensity or relationship to study drug or if new information becomes available, the SAE report must be updated and submitted within 24 hours to BMS (or designee) using the same procedure used for transmitting the initial SAE report. All SAEs must be followed to resolution or stabilization.

REPORTING OF SAES TO SPONSOR OR DESIGNEE

- SAEs, whether related or not related to study drug, and pregnancies must be reported to BMS (or designee) within 24 hours of awareness of the event.
- SAEs must be recorded on the SAE Report Form; pregnancies on a Pregnancy Surveillance Form (electronic or paper forms).
- The preferred method for SAE data reporting collection is through the eCRF.
- The paper SAE/pregnancy surveillance forms are only intended as a back-up option when the eCRF system is not functioning.
    - In this case, the paper forms are to be transmitted via email or confirmed facsimile (fax) transmission to:

SAE Email Address: Refer to Contact Information list.

SAE Facsimile Number: Refer to Contact Information list.

For studies capturing SAEs through electronic data capture (EDC), electronic submission is the required method for reporting. In the event the electronic system is unavailable for transmission, paper forms must be used and submitted immediately. When paper forms are used, the original paper forms are to remain on site.

SAE Telephone Contact (required for SAE and pregnancy reporting): Refer to Contact Information list

APPENDIX 4 WOMEN OF CHILDBEARING POTENTIAL DEFINITIONS AND METHODS OF CONTRACEPTION

DEFINITIONS

Woman of Childbearing Potential (WOCBP)

A woman is considered fertile following menarche and until becoming post-menopausal unless permanently sterile. Permanent sterilization methods include hysterectomy, bilateral salpingectomy, and bilateral oophorectomy.

Women in the following categories are not considered WOCBP

- Premenarchal
- Premenopausal female with 1 of the following:
  - Documented hysterectomy
  - Documented bilateral salpingectomy
  - Documented bilateral oophorectomy

Note: Documentation can come from the site personnel's review of the participant's medical records, medical examination, or medical history interview.

- Postmenopausal female
  - A postmenopausal state is defined as 12 months of amenorrhea in a woman over age 45 years in the absence of other biological or physiological causes. In addition, females under the age of 55 years must have a serum follicle stimulating hormone, (FSH) level > 40 mIU/mL to confirm menopause.

CONTRACEPTION GUIDANCE FOR FEMALE PARTICIPANTS OF CHILD BEARING POTENTIAL

One of the highly effective methods of contraception listed below is required during study duration and until the end of relevant systemic exposure, defined as 5 months after the end of study treatment.

Highly Effective Contraceptive Methods That Are User Dependent *Failure rate of <1% per year when used consistently and correctly.*[a]
• Combined (estrogen- and progestogen-containing) hormonal contraception associated with inhibition of ovulation[b] – oral – intravaginal – transdermal

• Progestogen-only hormonal contraception associated with inhibition of ovulation[b] – oral – injectable
Highly Effective Methods That Are User Independent
• Implantable progestogen-only hormonal contraception associated with inhibition of ovulation[b] • Intrauterine hormone-releasing system (IUS)[c] • Intrauterine device (IUD)[c] • Bilateral tubal occlusion
• Vasectomized partner *A vasectomized partner is a highly effective contraception method provided that the partner is the sole male sexual partner of the WOCBP and the absence of sperm has been confirmed. If not, an additional highly effective method of contraception should be used.*
• Sexual abstinence *Sexual abstinence is considered a highly effective method only if defined as refraining from heterosexual intercourse during the entire period of risk associated with the study drug. The reliability of sexual abstinence needs to be evaluated in relation to the duration of the study and the preferred and usual lifestyle of the participant.* • It is not necessary to use any other method of contraception when complete abstinence is elected. • WOCBP participants who choose complete abstinence must continue to have pregnancy tests, as specified in Section 2. • Acceptable alternate methods of highly effective contraception must be discussed in the event that the WOCBP participants chooses to forego complete abstinence
NOTES: [a] Typical use failure rates may differ from those when used consistently and correctly. Use should be consistent with local regulations regarding the use of contraceptive methods for participants participating in clinical studies. [b] Hormonal contraception may be susceptible to interaction with the study drug, which may reduce the efficacy of the contraceptive method. Hormonal contraception is permissible only when there is sufficient evidence that the IMP and other study medications will not alter hormonal exposures such that contraception would be ineffective or result in increased exposures that could be potentially hazardous. In this case, alternative methods of contraception should be utilized. [c] Intrauterine devices and intrauterine hormone releasing systems are acceptable methods of contraception in the absence of definitive drug interaction studies when hormone exposures from intrauterine devices do not alter contraception effectiveness

Unacceptable as a Sole Method of Contraception
• Male or female condom with or without spermicide. Male and female condoms cannot be used simultaneously

- Diaphragm with spermicide
- Cervical cap with spermicide
- Vaginal Sponge with spermicide
- Progestogen-only oral hormonal contraception, where inhibition of ovulation is not the primary mechanism of action
- Periodic abstinence (calendar, symptothermal, post-ovulation methods)
- Withdrawal (coitus interruptus).
- Spermicide only
- Lactation amenorrhea method (LAM)

*** Local laws and regulations may require use of alternative and/or additional contraception methods.**

CONTRACEPTION GUIDANCE FOR MALE PARTICIPANTS WITH PARTNER(S) OF CHILD BEARING POTENTIAL.

Male participants with female partners of childbearing potential are eligible to participate if they agree to the following during the treatment and until the end of relevant systemic exposure.

- Inform any and all partner(s) of their participation in a clinical drug study and the need to comply with contraception instructions as directed by the investigator.
- Male participants are required to use a condom for study duration and until end of relevant systemic exposure defined as 7 months after the end of study treatment.
- Female partners of males participating in the study to consider use of effective methods of contraception until the end of relevant systemic exposure, defined as 7 months after the end of treatment in the male participant.
- Male participants with a pregnant or breastfeeding partner must agree to remain abstinent from penile vaginal intercourse or use a male condom during each episode of penile penetration during the treatment and until 7 months after the end of study treatment.
- Refrain from donating sperm for the duration of the study treatment and until 7 months after the end of study treatment.

COLLECTION OF PREGNANCY INFORMATION

Guidance for collection of Pregnancy Information and outcome of pregnancy on the Pregnancy Surveillance Form is provided in Section 9.2.5 and the Appendix for Adverse Events and Serious Adverse Events Definitions and procedures for Evaluating, Follow-up and Reporting

APPENDIX 5 MANAGEMENT ALGORITHMS

These general guidelines constitute guidance to the Investigator and may be supplemented by discussions with the Medical Monitor representing the Sponsor. The guidance applies to all immuno-oncology agents and regimens.

A general principle is that differential diagnoses should be diligently evaluated according to standard medical practice. Non-inflammatory etiologies should be considered and appropriately treated.

Corticosteroids are a primary therapy for immuno-oncology drug-related adverse events. The oral equivalent of the recommended IV doses may be considered for ambulatory patients with low-grade toxicity. The lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.

Consultation with a medical or surgical specialist, especially prior to an invasive diagnostic or therapeutic procedure, is recommended.

The frequency and severity of the related adverse events covered by these algorithms will depend on the immuno-oncology agent or regimen being used.

GI Adverse Event Management Algorithm

Rule out non-inflammatory causes. If non-inflammatory cause is identified, treat accordingly and continue I-O therapy. Opiates/narcotics may mask symptoms of perforation. Infliximab should not be used in cases of perforation or sepsis.

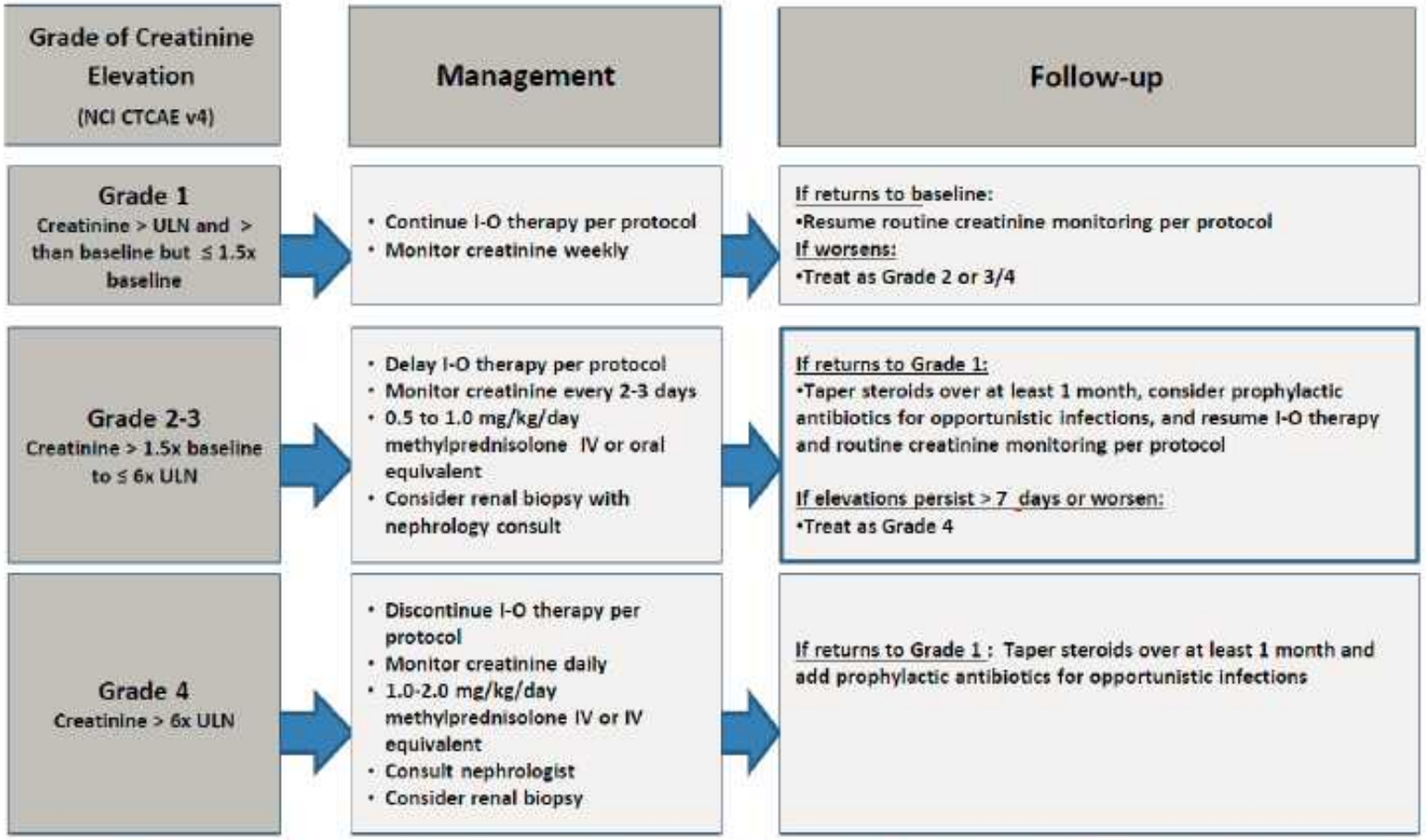
Grade of Diarrhea/ Colitis (NCI CTCAE v4)	Management	Follow-up
Grade 1 Diarrhea: < 4 stools/day over baseline; Colitis: asymptomatic	• Continue I-O therapy per protocol • Symptomatic treatment	• Close monitoring for worsening symptoms. • Educate patient to report worsening immediately If worsens: • Treat as Grade 2 or 3/4
Grade 2 Diarrhea: 4-6 stools per day over baseline; IV fluids indicated <24 hrs; not interfering with ADL Colitis: abdominal pain; blood in stool	• Delay I-O therapy per protocol • Symptomatic treatment	If improves to grade 1: • Resume I-O therapy per protocol If persists > 5-7 days or recur: • 0.5-1.0 mg/kg/day methylprednisolone or oral equivalent • When symptoms improve to grade 1, taper steroids over at least 1 month, consider prophylactic antibiotics for opportunistic infections, and resume I-O therapy per protocol. If worsens or persists > 3-5 days with oral steroids: • Treat as grade 3/4
Grade 3-4 Diarrhea (G3): ≥7 stools per day over baseline; incontinence; IV fluids ≥24 hrs; interfering with ADL Colitis (G3): severe abdominal pain, medical intervention indicated, peritoneal signs G4: life-threatening, perforation	• Discontinue I-O therapy per protocol • 1.0 to 2.0 mg/kg/day methylprednisolone IV or IV equivalent • Add prophylactic antibiotics for opportunistic infections • Consider lower endoscopy	If improves: • Continue steroids until grade 1, then taper over at least 1 month If persists > 3-5 days, or recurs after improvement: • Add infliximab 5 mg/kg (if no contraindication). Note: Infliximab should not be used in cases of perforation or sepsis

Patients on IV steroids may be switched to an equivalent dose of oral corticosteroids (e.g. prednisone) at start of tapering or earlier, once sustained clinical improvement is observed. Lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.

27-Jun-2018

Renal Adverse Event Management Algorithm

Rule out non-inflammatory causes. If non-inflammatory cause, treat accordingly and continue I-O therapy.

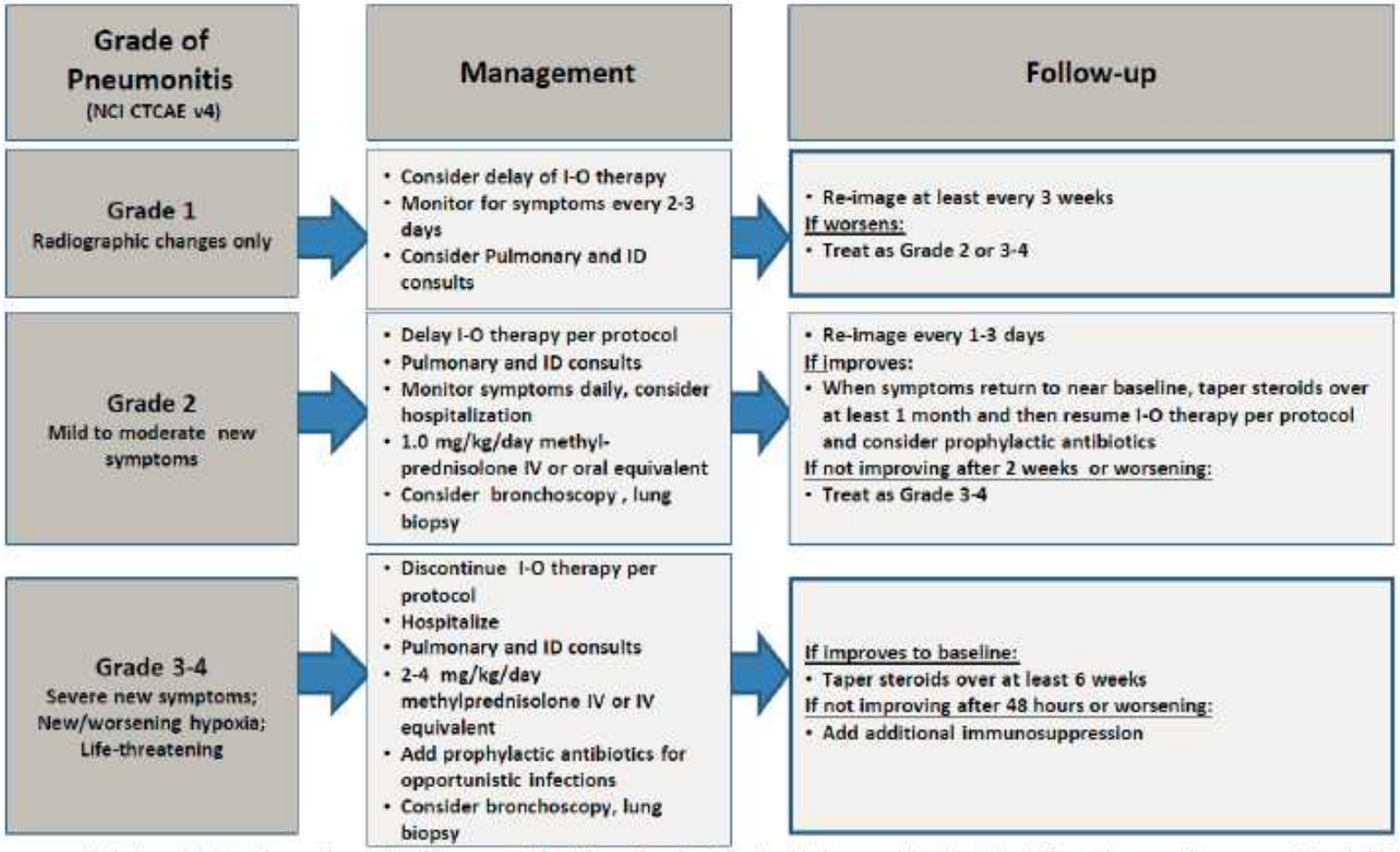


Patients on IV steroids may be switched to an equivalent dose of oral corticosteroids (e.g. prednisone) at start of tapering or earlier, once sustained clinical improvement is observed. Lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.

27-Jun-2018

Pulmonary Adverse Event Management Algorithm

Rule out non-inflammatory causes. If non-inflammatory cause, treat accordingly and continue I-O therapy. Evaluate with imaging and pulmonary consultation.

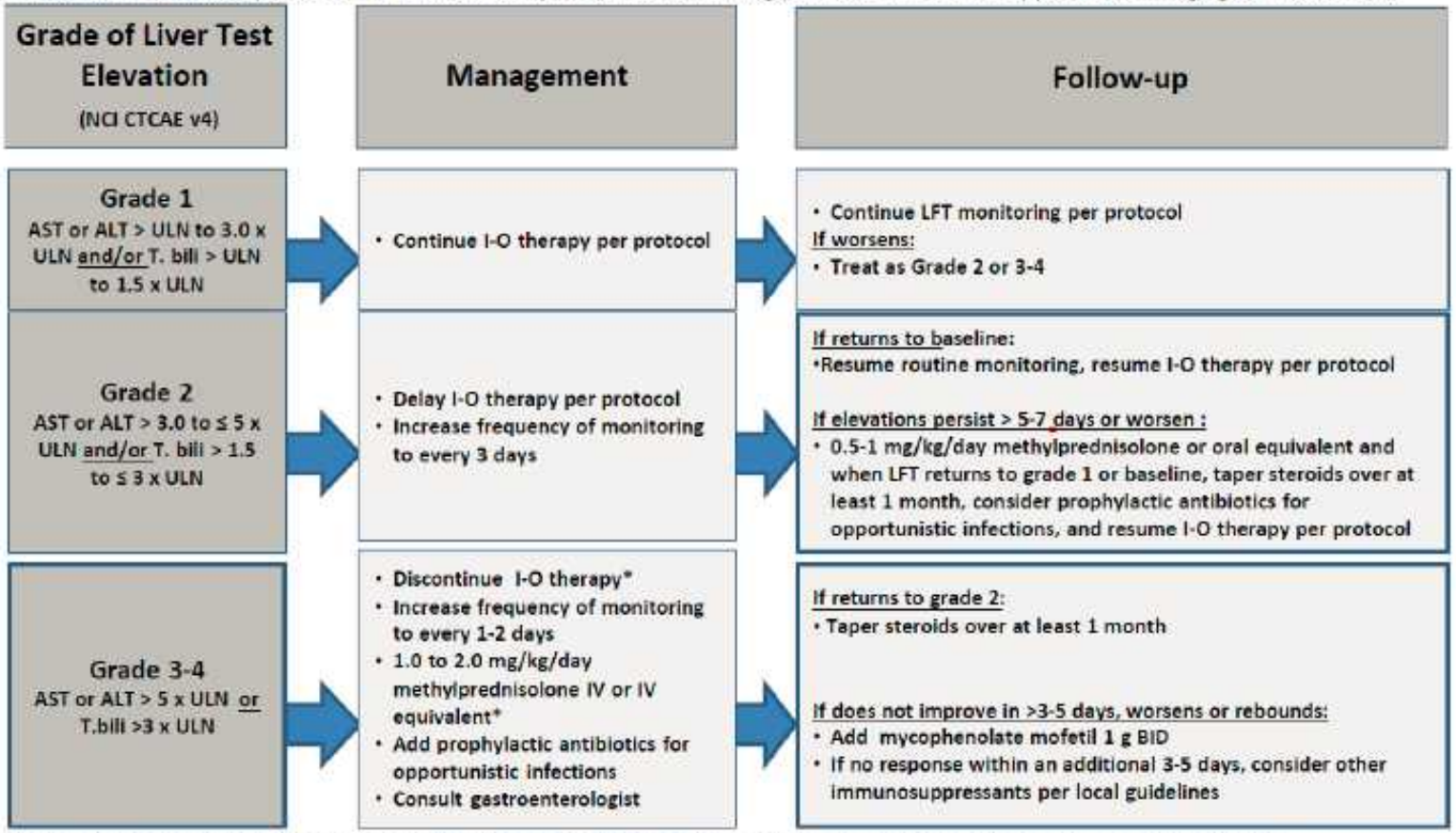


Patients on IV steroids may be switched to an equivalent dose of oral corticosteroids (e.g. prednisone) at start of tapering or earlier, once sustained clinical improvement is observed. Lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids

27-Jun-2018

Hepatic Adverse Event Management Algorithm

Rule out non-inflammatory causes. If non-inflammatory cause, treat accordingly and continue I-O therapy. Consider imaging for obstruction.



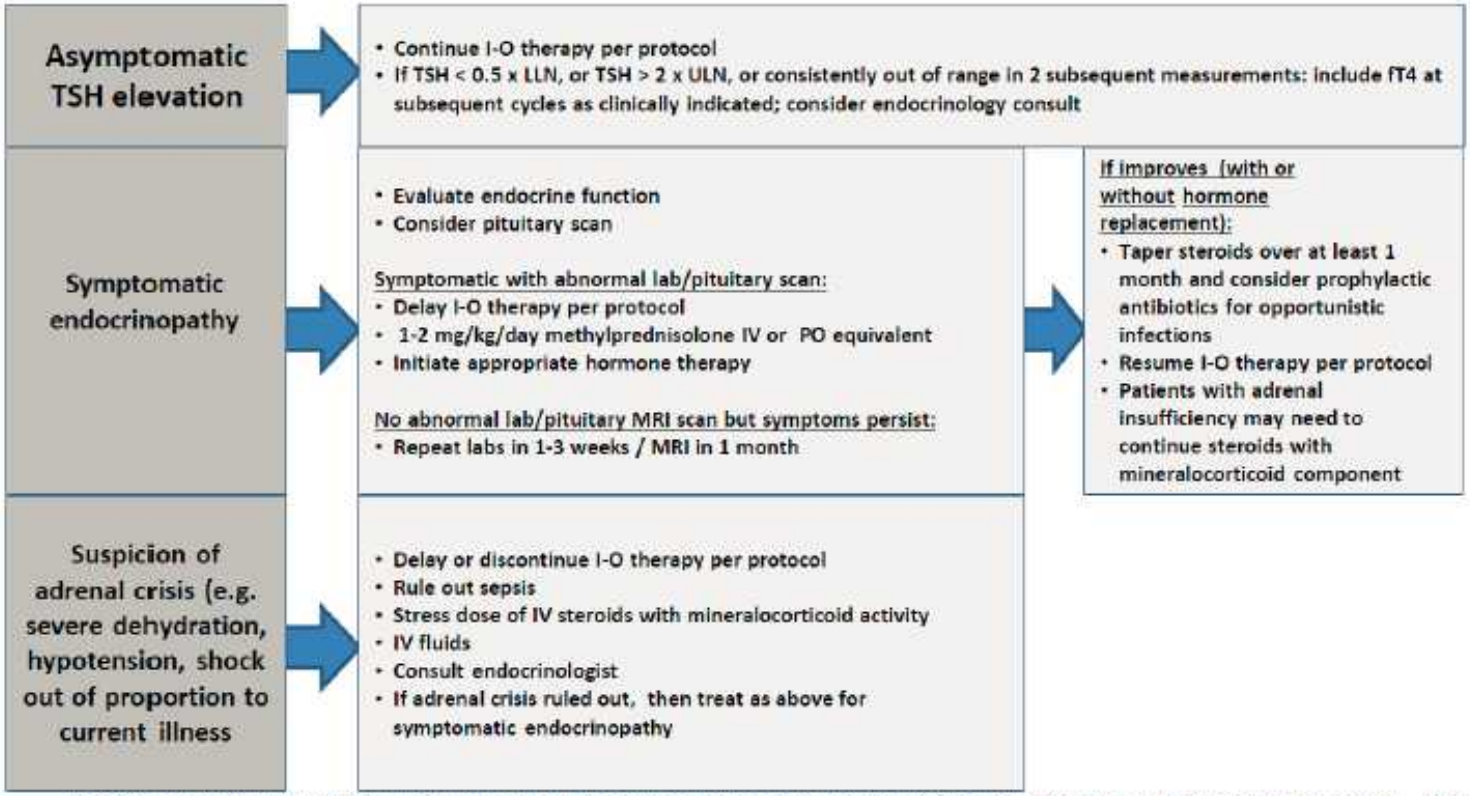
Patients on IV steroids may be switched to an equivalent dose of oral corticosteroids (e.g. prednisone) at start of tapering or earlier, once sustained clinical improvement is observed. Lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.

*The recommended starting dose for grade 4 hepatitis is 2 mg/kg/day methylprednisolone IV.

27-Jun-2018

Endocrinopathy Adverse Event Management Algorithm

Rule out non-inflammatory causes. If non-inflammatory cause, treat accordingly and continue I-O therapy. Consider visual field testing, endocrinology consultation, and imaging.

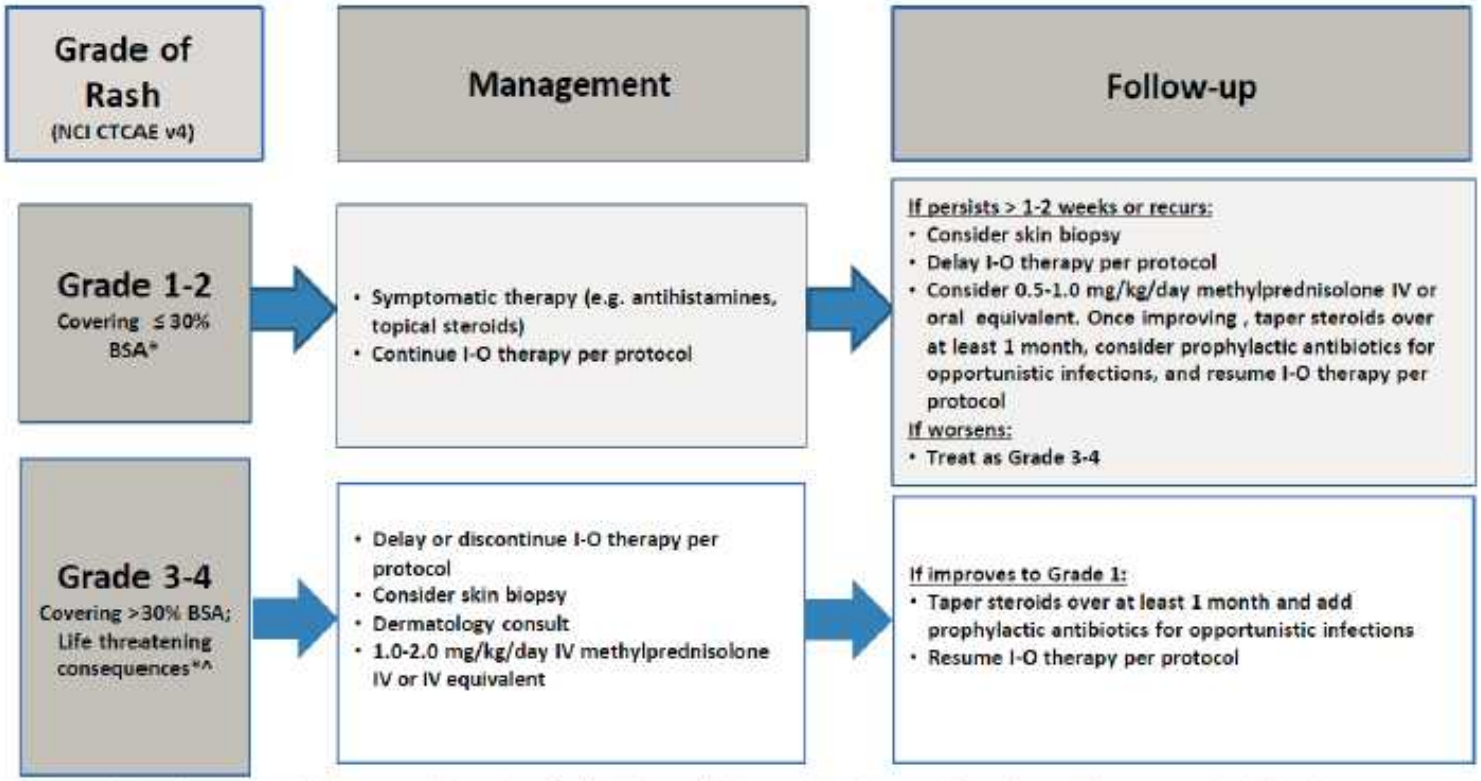


Patients on IV steroids may be switched to an equivalent dose of oral corticosteroids (e.g. prednisone) at start of tapering or earlier, once sustained clinical improvement is observed. Lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.

27-Jun-2018

Skin Adverse Event Management Algorithm

Rule out non-inflammatory causes. If non-inflammatory cause, treat accordingly and continue I-O therapy.



Patients on IV steroids may be switched to an equivalent dose of oral corticosteroids (e.g. prednisone) at start of tapering or earlier, once sustained clinical improvement is observed. Lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.

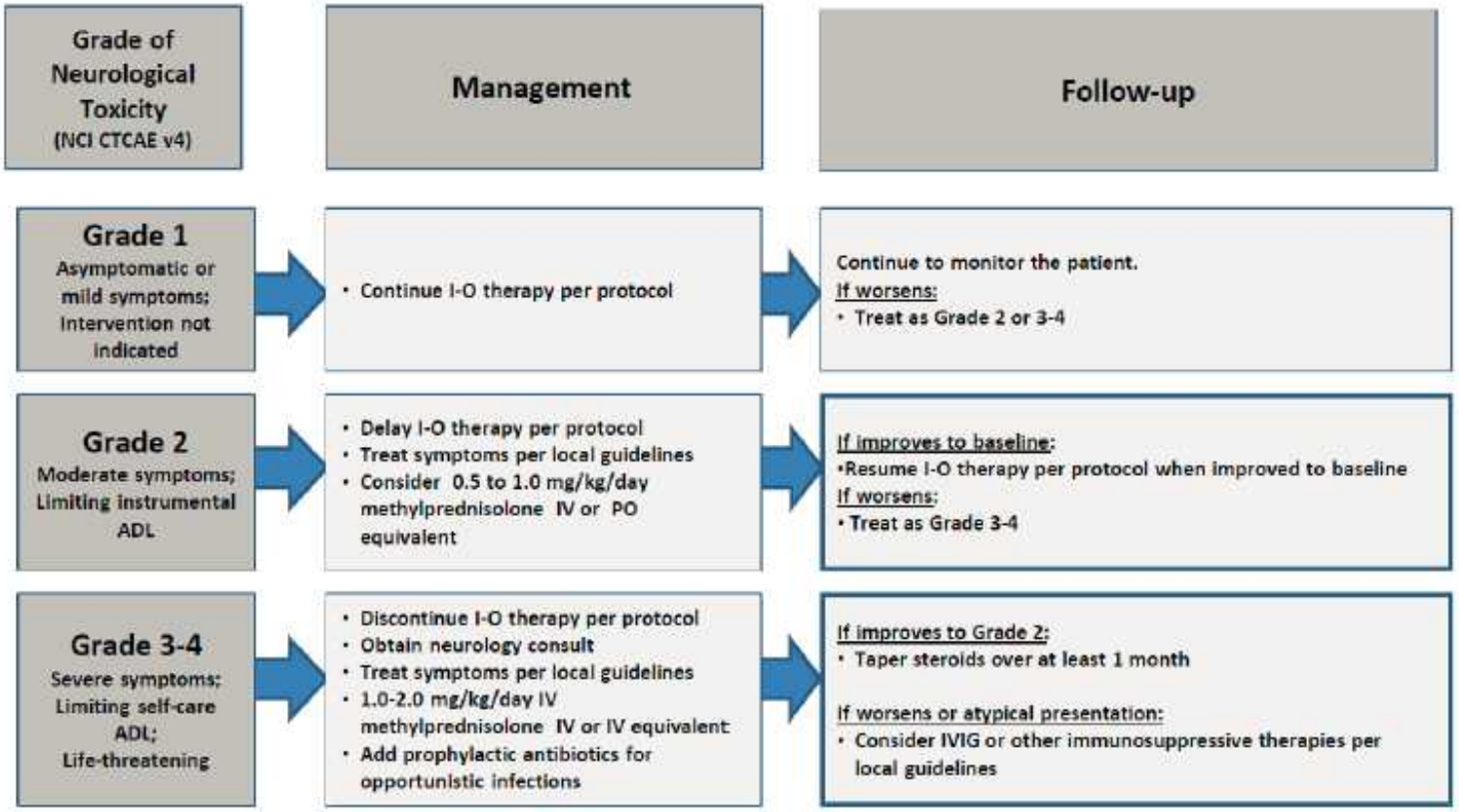
*Refer to NCI CTCAE v4 for term-specific grading criteria.

^If SJS/TEN is suspected, withhold I-O therapy and refer patient for specialized care for assessment and treatment. If SJS or TEN is diagnosed, permanently discontinue I-O therapy.

27-Jun-2018

Neurological Adverse Event Management Algorithm

Rule out non-inflammatory causes. If non-inflammatory cause, treat accordingly and continue I-O therapy.



Patients on IV steroids may be switched to an equivalent dose of oral corticosteroids (e.g. prednisone) at start of tapering or earlier, once sustained clinical improvement is observed. Lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.

27-Jun-2018

APPENDIX 6 INTERNATIONAL METASTATIC RCC DATABASE CONSORTIUM (IMDC) PROGNOSTIC CRITERIA

Adverse Prognostic Factors
Clinical
KPS < 80% Time from initial diagnosis (including original localized disease if applicable) to treatment < 1 year
Laboratory
Hemoglobin < LLN Corrected calcium > 10 mg/dL Absolute neutrophil count > ULN Platelet count > ULN

Note: The corrected calcium criterion was adapted from to Heng et al, 2009 to account for local laboratories that may not provide an ULN for corrected calcium.

Abbreviations: KPS= Karnofsky Performance Status; LLN = Lower limit of normal; ULN = Upper limit of normal

Corrected calcium (mg/dL) = measured total Ca (mg/dL) + 0.8 (4.0 - serum albumin [g/dL]), where 4.0 represents the average albumin level in g/dL.

Corrected calcium (mmol/L) = measured total Ca (mmol/L) + 0.02 (40 - serum albumin [g/L]), where 40 represents the average albumin level in g/L

Risk Group Based on Number of Adverse Prognostic Factors	
Number of Adverse Prognostic Factors Present	**Risk Group**
0	Favorable
1-2	Intermediate
3-6	Poor

Reference: Heng D, Xie W, Regan M, et al. Prognostic factors for overall survival in patients with metastatic renal cell carcinoma treated with vascular endothelial growth factor-targeted agents: results from a large, multicenter study. J Clin Oncol 2009; 27(34):5794-5799.

APPENDIX 7 PERFORMANCE STATUS SCORES

	SCALES		
STATUS	**KARNOFSKY**	**ZUBROD-ECOG-WHO**	**STATUS**
Normal, no complaints	100	0	Normal activity
Able to carry on normal activities Minor signs or symptoms of disease	90	0	
Normal activity with effort	80	1	Symptoms, but fully ambulatory
Cares for self. Unable to carry on normal activity or to do active work	70	1	
Requires occasional assistance, but able to care for most of his needs	60	2	Symptomatic, but in bed < 50% of the day.
Requires considerable assistance and frequent medical care	50	2	
Disabled. Requires special care and assistance	40	3	Needs to be in bed > 50% of the day, but not bedridden
Severely disabled. Hospitalization indicated though death non imminent	30	3	
Very sick. Hospitalization necessary. Active supportive treatment necessary	20	4	Unable to get out of bed
Moribund	10	4	
Dead	0	5	Dead

APPENDIX 8 RADIOLOGIC EVALUATION CRITERIA IN SOLID TUMOURS VERSION 1.1 (RECIST CRITERIA 1.1)

1 ASSESSMENT OF OVERALL TUMOR BURDEN AND MEASURABLE DISEASE

Subjects must have measureable disease to be eligible for this study.

To assess objective response or future progression, it is necessary to estimate the *overall tumor burden at baseline* and use this as a comparator for subsequent measurements. Measurable disease is defined by the presence of at least 1 measurable tumor lesion. When computed tomography (CT) scans have slice thickness greater than 5 mm, the minimum size for a measurable lesion should be twice the slice thickness.

At baseline, tumor lesions/lymph nodes will be categorized measurable or nonmeasurable, which are discussed below.

1.1 Measurable Lesions

Measurable lesions must be accurately measured in at least 1 dimension (longest diameter in the plane of the measurement to be recorded) with a minimum size of:

- 10 mm by CT/magnetic resonance imaging (MRI) scan - (CT/MRI scan slice thickness no greater than 5 mm)
- 10 mm caliper measurement by clinical exam (lesions which cannot be accurately measured with calipers should be recorded as nonmeasurable)
- 20 mm by chest x-ray
- Malignant lymph nodes: To be considered pathologically enlarged and measurable, a lymph node must be ≥ 15 mm in short axis when assessed by CT scan (CT scan slice thickness recommended to be no greater than 5 mm). At baseline and in follow-up, only the short axis will be measured and followed.

1.2 Non-measurable Lesions

- All other lesions, including small lesions (longest diameter < 10 mm or pathological lymph nodes with ≥ 10 to < 15 mm short axis), as well as truly non-measurable lesions
- Lesions considered truly nonmeasurable include: leptomeningeal disease, ascites, pleural or pericardial effusion, inflammatory breast disease, lymphangitic involvement of skin or lung, and abdominal masses/abdominal organomegaly identified by physical exam that is not measurable by reproducible imaging techniques.

1.3 Special Considerations Regarding Lesion Measurability

1.3.1 Bone Lesions

- Bone scan, positron emission tomography (PET) scan, or plain films are not considered adequate imaging techniques to measure bone lesions. However, these techniques can be used to confirm the presence or disappearance of bone lesions.
- Lytic bone lesions or mixed lytic-blastic lesions, with identifiable soft tissue components that can be evaluated by cross sectional imaging techniques such as CT or MRI, can be considered

measurable lesions if the soft tissue component meets the definition of measurability described above.

- Blastic bone lesions are nonmeasurable.

1.3.2 Cystic Lesions

- Lesions that meet the criteria for radiographically defined simple cysts should not be considered as malignant lesions (neither measurable nor nonmeasurable) since they are, by definition, simple cysts.
- 'Cystic lesions' thought to represent cystic metastases can be considered as measurable lesions if they meet the definition of measurability described above. However, if non-cystic lesions are present in the same subject, these are preferred for selection as target lesions.

1.3.3 Lesions with Prior Local Treatment

Tumor lesions situated in a previously irradiated area or in an area subjected to other locoregional therapy are usually not considered measurable, unless there has been demonstrated progression in the lesion. Measurable lesions may be in an irradiated field as long as there is documented progression, and the lesion(s) can be reproducibly measured.

1.4 Specifications by Methods of Measurements

1.4.1 Measurement of Lesions

All measurements should be recorded in metric notation (mm). All baseline evaluations should be performed as close as possible to the treatment start and never more than 30 days before the beginning of the treatment.

1.4.2 Method of Assessment

The **same method of assessment and the same technique should be used** to characterize each identified and reported lesion at baseline and during follow-up. Imaging-based evaluation should always be done rather than clinical examination, unless the lesion(s) being followed cannot be imaged but are assessable by clinical exam.

1.4.2.1 CT/MRI Scan

CT/MRI is the best currently available and reproducible method to measure lesions selected for response assessment. Measurability of lesions on CT/MRI scan is based on the assumption that CT/MRI slice thickness is 5 mm or less. When CT scans have slice thickness greater than 5 mm, the minimum size for a measurable lesion should be twice the slice thickness.

1.4.2.2 Chest X-ray

Chest CT is preferred over chest X-ray, particularly when progression is an important endpoint, since CT is more sensitive than X-ray in identifying new lesions. However, lesions on chest X-ray may be considered measurable if they are clearly defined and surrounded by aerated lung.

1.4.2.3 Clinical Lesions

Clinical lesions will only be considered measurable when they are superficial and ≥ 10 mm diameter as assessed using calipers. For the case of skin lesions, documentation by color

photography including a ruler to estimate the size of the lesion is suggested. As previously noted, when lesions can be evaluated by both clinical examination and imaging, imaging evaluation should be undertaken since it is more objective and may also be reviewed at the end of the study.

1.4.2.4 Ultrasound

Ultrasound is **not** useful in assessment of lesion size and should not be used as a method of measurement. If new lesions are identified by ultrasound in the course of the study, confirmation by CT or MRI is advised.

1.4.2.5 Endoscopy, Laparoscopy

The utilization of these techniques for objective tumor evaluation is **not** advised.

1.4.2.6 Tumor Markers

Tumor markers **alone** cannot be used to assess objective tumor response.

2 BASELINE DOCUMENTATION OF 'TARGET' AND 'NON-TARGET' LESIONS

2.1 Target Lesions

When more than 1 measurable lesion is present at baseline, all lesions up to **a maximum of 5 lesions total (and a maximum of 2 lesions per organ) representative of all involved organs should be identified as *target lesions*** and will be recorded and measured at baseline.

Target lesions should be selected on the basis of their **size** (lesions with the longest diameter), be representative of all involved organs, and should lend themselves to **reproducible repeated measurements.**

A **sum of the diameters** (longest for non-nodal lesions, short axis for nodal lesions) for all target lesions will be calculated and reported as the **baseline sum diameters.** If lymph nodes are to be included in the sum, then as noted below, only the **short** axis is added into the sum. The baseline sum diameters will be used as reference to further characterize any objective tumor regression in the measurable dimension of the disease.

2.1.1 Lymph Nodes

Lymph nodes merit special mention since they are normal anatomical structures which may be visible by imaging even if not involved by tumor. Pathological nodes, which are defined as measurable and may be identified as target lesions, must meet the criterion of a **short axis of ≥ 15 mm by CT scan**. Only the short axis of these nodes will contribute to the baseline sum. Nodes that have a short axis < 10 mm are considered nonpathological and should not be recorded or followed.

2.2 Non-target Lesions

All other lesions (or sites of disease), including pathological lymph nodes, should be identified as **non-target lesions** and should also be recorded at baseline. Measurements are not required, and these lesions should be followed as **'present', 'absent', or in rare cases, 'unequivocal progression'.** In addition, it is possible to record multiple non-target lesions involving the same

organ as a single item on the case record form (e.g., 'multiple enlarged pelvic lymph nodes' or 'multiple liver metastases').

3 TUMOR RESPONSE EVALUATION

3.1 Evaluation of Target Lesions

- Complete Response (CR): **Disappearance of all target lesions.** Any pathological lymph nodes (whether target or nontarget) must have reduction in short axis to < 10 mm.
- Partial Response (PR): At least a **30% decrease in the sum of diameters of target lesions,** taking as reference the baseline sum diameters.
- Progressive Disease (PD): At least a **20% increase in the sum of diameters of target lesions,** taking as reference the **smallest sum on study** (this includes the baseline sum if that is the smallest on study). In addition to the relative increase of 20%, the sum must also demonstrate an **absolute increase of at least 5 mm**. (Note*:* The appearance of 1 or more new lesions is also considered progression).
- Stable Disease (SD): Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD, taking as reference the smallest sum on study.

3.1.1 Special Notes on the Assessment of Target Lesions

3.1.1.1 Lymph Nodes

Lymph nodes merit special mention since they are normal anatomical structures which may be visible by imaging even if not involved by tumor. Pathological nodes, which are defined as measurable and may be identified as target lesions, must meet the criterion of a **short axis of ≥ 15 mm by CT scan**. Only the short axis of these nodes will contribute to the baseline sum. Nodes that have a short axis <10 mm are considered nonpathological and should not be recorded or followed.

3.1.1.2 Target Lesions That Become 'Too Small to Measure'

All lesions (nodal and nonnodal) recorded at baseline should have their actual measurements recorded at each subsequent evaluation, even when very small (e.g., 2 mm). If the radiologist is able to provide an actual measurement, that should be recorded even if it is below 5 mm.

However, when such a lesion becomes difficult to assign an exact measure to then:

- If it is the opinion of the radiologist that the lesion has likely disappeared, the measurement should be recorded as 0 mm.
- If the lesion is believed to be present and is faintly seen but too small to measure, a default value of 5 mm should be assigned (Note: In case of a lymph node believed to be present and faintly seen but too small to measure, a default value of 5 mm should be assigned in this circumstance as well). This default value is derived from the 5 mm CT slice thickness (but should not be changed with varying CT slice thickness).

3.1.1.3 Target Lesions that Split or Coalesce on Treatment

- When non-nodal lesions fragment, the longest diameters of the fragmented portions should be added together to calculate the target lesion sum.

- As lesions coalesce, a plane between them may be maintained that would aid in obtaining maximal diameter measurements of each individual lesion. If the lesions have truly coalesced such that they are no longer separable, the vector of the longest diameter in this instance should be the maximal longest diameter for the 'coalesced lesion'

3.2 Evaluation of Non-target Lesions

While some non-target lesions may actually be measurable, they need not be measured and instead should be assessed only qualitatively at the time points specified in the protocol.

- CR: Disappearance of all non-target lesions. All lymph nodes must be nonpathological in size (< 10 mm short axis).
- PD: Unequivocal progression of existing non-target lesions. (Note: The appearance of 1 or more new lesions is also considered progression).
- NonCR/NonPD: Persistence of 1 or more non-target lesion(s).

3.2.1 Special Notes on Assessment of Non-target Lesions

The concept of progression of non-target disease requires additional explanation as discussed below.

3.2.1.1 When the Subject Also has Measurable Disease

- To achieve unequivocal progression on the basis of the non-target disease, there must be an overall level of substantial worsening in non-target disease, such that even in presence of SD or PR in target disease, the overall tumor burden has increased sufficiently to merit discontinuation of therapy.
- A modest 'increase' in the size of 1 or more non-target lesions is usually not sufficient to quality for unequivocal progression status.

3.2.1.2 When the Subject has Only Non-measurable Disease

- To achieve unequivocal progression on the basis of the non-target disease, there must be an overall level of substantial worsening, such that the overall tumor burden has increased sufficiently to merit discontinuation of therapy.
- A modest increase in the size of 1 or more non-target lesions is usually not sufficient to qualify for unequivocal progression status.
- Because worsening in non-target disease cannot be easily quantified (by definition, if all lesions are nonmeasurable), a useful test that can be applied when assessing subjects for unequivocal progression is to consider if the increase in overall disease burden based on the change in non-measurable disease is comparable in magnitude to the increase that would be required to declare PD for measurable disease, ie, an increase in tumor burden representing an additional 73% increase in volume (which is equivalent to a 20% increase diameter in a measurable lesion). Examples include an increase in a pleural effusion from 'trace' to 'large', an increase in lymphangitic disease from localized to widespread, or may be described in protocols as 'sufficient to require a change in therapy'.
- If unequivocal progression is seen, the subject should be considered to have had overall PD at that point.

3.2.1.3 Tumor Markers

Tumor markers will not be used to assess objective tumor responses.

3.3 New Lesions

The appearance of new malignant lesions denotes disease progression. The finding of a new lesion should be unequivocal, that is, not attributable to differences in scanning technique, change in imaging modality, or findings thought to represent something other than tumor (for example, some new bone lesions may be simply healing or flare of preexisting lesions). This is particularly important when the subject's baseline lesions show PR or CR. For example, necrosis of a liver lesion may be reported on a CT scan report as a new cystic lesion, which it is not.

A lesion identified on a follow-up study in an anatomical location that was **not** scanned at baseline is considered a new lesion and will indicate disease progression. An example of this is the subject who has visceral disease at baseline and while on study, has a CT or MRI brain scan ordered that reveals metastases. The subject's brain metastases are considered to be evidence of PD even if he/she did not have brain imaging at baseline.

If a new lesion is equivocal, for example because of its small size, continued therapy and follow-up evaluation will clarify if it represents truly new disease. If repeat scans confirm there is definitely a new lesion, then progression should be declared using the date of the initial scan.

4 RESPONSE CRITERIA

4.1 Time Point Response

A response assessment should occur at each time point specified in the protocol.

For subjects who have **measurable disease** at baseline, Table 1 provides a summary of the overall response status calculation at each time point.

Table 1: Subjects with Target (+/– Non-target) Disease

Target Lesions	Non-target Lesions	New Lesions	Overall Response
CR	CR	No	CR
CR	NonCR/nonPD	No	PR
CR	Not evaluated	No	PR
PR	NonPD or not all evaluated	No	PR
SD	NonPD or not all evaluated	No	SD
Not all evaluated	NonPD	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD

Abbreviations: CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, NE = not evaluable.

4.1.1 Missing Assessments and Not Evaluable Designation

When no imaging/measurement is done at a particular time point, the subject is **not evaluable (NE)** at that time point. If only a subset of lesion measurements are made at an assessment, the case is also considered NE at that time point, unless a convincing argument can be made that the contribution of the individual missing lesion(s) would not have changed the assigned time point response.

4.2 Best Overall Response: All Time Points

The best overall response is determined once all the data for the subject is known. It is the best response recorded from the start of the study treatment to the date of radiographic progression per RECIST 1.1 or the date of subsequent anti-cancer therapy (including tumor-directed radiotherapy and tumor-directed surgery), whichever occurs first, taking into account any requirement for confirmation. The subject's best overall response assignment will depend on the findings of both target and non-target disease and will also take into consideration the appearance of new lesions.

Complete or partial responses may be claimed only if the criteria for each are met at a subsequent time point of ≥ 4 weeks later. In this circumstance, the best overall response can be interpreted as specified in Table 2.

When SD is believed to be best response, it must meet the protocol specified minimum time from baseline. In this protocol, measurements must have met the SD criteria at least once after study entry at a minimum interval of 6 weeks from randomization in order for SD to be the best response.

Table 2: Best Overall Response When Confirmation of CR and PR is Required

Overall Response	Overall Response	Best Overall Response
First Time Point	Subsequent Time Point	
CR	CR	CR
CR	PR	SD, PD, or PR[a]
CR	SD	SD provided minimum criteria for SD duration met, otherwise, PD
CR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
CR	NE	SD provided minimum criteria for SD duration met, otherwise NE
PR	CR	PR
PR	PR	PR
PR	SD	SD
PR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
PR	NE	SD provided minimum criteria for SD duration met, otherwise NE

Table 2: Best Overall Response When Confirmation of CR and PR is Required

Overall Response	Overall Response	Best Overall Response
NE	NE	NE

[a] Abbreviations: CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, NE = not evaluable.

If a CR is truly met at first time point, then any disease seen at a subsequent time point, even disease meeting PR criteria relative to baseline, makes the disease PD at that point (since disease must have reappeared after CR). Best response would depend on whether minimum duration for SD was met. However, sometimes 'CR' may be claimed when subsequent scans suggest small lesions were likely still present and in fact the subject had PR, not CR at the first time point. Under these circumstances, the original CR should be changed to PR and the best response is PR.

4.2.1 Confirmation Scans

Verification of Response: To be assigned a status of CR or PR, changes in tumor measurements must be confirmed by consecutive repeat assessments that should be performed no less than 28 days after the criteria for response are first met. For this study, the next scheduled tumor assessment can meet this requirement.

4.3 Duration of Response

4.3.1 Duration of Overall Response

The duration of overall response is measured from the time measurement criteria are first met for CR/PR (whichever is first recorded) until the first date that progressive disease is objectively documented (taking as reference for PD the smallest measurements recorded on study), subsequent anti-cancer therapy (including tumor-directed radiotherapy and tumor-directed surgery) is initiated, or the participant dies, whichever occurs first.

The duration of overall complete response is measured from the time measurement criteria are first met for CR until the first date that progressive disease is objectively documented, subsequent anti-cancer therapy (including tumor-directed radiotherapy and tumor-directed surgery) is initiated, or the participant dies, whichever occurs first.

4.3.2 Duration of Stable Disease

If SD is the best overall response, the duration of SD is measured from the date of randomization until the criteria for progression are met, taking as reference the smallest sum on study (if the baseline sum is the smallest, this is the reference for calculation of PD), subsequent anti-cancer therapy (including tumor-directed radiotherapy and tumor-directed surgery) is initiated, or the participant dies, whichever occurs first..

Reference: Eisenhauer EA, Therasse P, Bogaerts J, et al. New Response Evaluation Criteria in Solid Tumours: Revised RECIST Guideline (Version 1.1). Eur J Cancer 2009;45:228-247.

APPENDIX 9 INDUCERS AND INHIBITORS OF CYP3A4

Strong Inducers and Inhibitors of CYP3A4
CYP3A4 Inducers
Phenytoin
Carbamazepine
Rifampin
Rifabutin
Rifapentine
Phenobarbital
Dexamethasone
CYP3A4 Inhibitors
Ketoconazole
Itraconazole
Voriconazole
Clarithromycin
Erythromycin
Telithromycin
Nefazodone
Saquinavir
Ritonavir
Atazanavir
Indinavir
Nelfinavir

Notes: The above list is not exhaustive. Please refer to the drug interaction tables at the link below for lists of substrates, inducers, and inhibitors of selected CYP450 isozyme pathways. (Http://medicine.iupui.edu/clinpharm/ddis/table.aspx).

Grapefruit, grapefruit juice and other foods that are known to inhibit CYP3A4 activity should be avoided during treatment.

St. John's Wort (*Hypericum perforatum*) is known to be an inducer of CYP3A4 and should be avoided during treatment.

APPENDIX 10 SUPPORTIVE CARE GUIDELINES FOR THE MANAGEMENT OF HAND FOOT SYNDROME (HFS)

Supportive Care Guidelines for the Management of Hand Foot Syndrome (HFS)

Management of HFS can begin before any symptoms occur. Several prophylactic measures may be taken to prevent or reduce the severity of HFS. Before therapy with begins, a full-body skin exam should be performed, with a special emphasis on hyperkeratotic areas on palms and soles and any deformities. Patients can receive a pedicure, using properly sterilized utensils, to remove any preexisting hyperkeratotic areas or calluses that may predispose them to developing HFS reaction. Patients should be advised to reduce the exposure of their hands and feet to hot water, either through dishwashing or hot baths and showers, because this is believed to exacerbate symptoms, and patients frequently report symptomatic relief with cold water.

Before initiating treatment:

- Check condition of hands and feet
- Suggest a manicure/pedicure, when indicated
- Recommend pumice stone use for callus or 'rough spot' removal

During treatment:

- Avoid pressure points
- Avoid items that rub, pinch, or create friction
- Apply non-urea based skin-hydrating creams liberally
- Keratolytic creams: Use sparingly and only to affected (hyperkeratotic) areas.

Urea-based creams:

- Salicylic acid 6%
- Alpha hydroxy acid (AHA) based creams
- Concentrations of approximately 5-8% provide gentle chemical exfoliation
- Apply liberally two times each day

Topical analgesics like lidocaine 2% should be considered for pain control

Topical corticosteroids like clobetasol 0.05% should be considered for patients with grade 2 or 3 hand-foot skin reaction. Avoid systemic steroids.

Cushions:

- Protect tender areas
- Use socks/gloves to cover moisturizing creams
- Wear well-padded footwear
- Use insole cushions or inserts (e.g., silicon, gel)
- Foot soaks with tepid water and Epsom salts

APPENDIX 11 SUPPORTIVE CARE GUIDELINES FOR THE MANAGEMENT OF DIARRHEA FOR CABOZANTINIB

Participants should be instructed to notify their physician immediately at the first signs of poorly formed or loose stool or an increased frequency of bowel movements. Guidelines for the evaluation and management of diarrhea are shown in Table 1.

Administration of antidiarrheal/antimotility agents is recommended at the first sign of diarrhea as initial management. Some participants may require concomitant treatment with more than 1 antidiarrheal agent. When therapy with antidiarrheal agent does not control diarrhea to tolerable levels, cabozantinib should be temporarily interrupted or dose reduced. When the diarrhea is controlled, retreatment with cabozantinib may be acceptable per investigator decision.

In addition, general supportive measures should be implemented such as continuous oral isotonic hydration, correction of fluid and electrolyte abnormalities, small frequent meals, and stopping lactose-containing products, high-fat meals, and alcohol.

Recurrent or prolonged diarrhea can be associated with anal or perianal skin erosions which increase the risk for anal abscesses, fistulas, or proctitis. Good hygiene should be emphasized. Regular examinations of the perianal region should be performed wherever diarrhea has occurred during treatment with cabozantinib. Infections of the perianal region should be treated per local guidelines.

Table 1: Guidelines for Management of Treatment-Emergent Diarrhea for Cabozantinib

Status	Management
Tolerable Grade 1-2 (duration < 48 h)	• Continue with study treatment and consider dose reduction • Initiate treatment with an antidiarrheal agent (e.g., loperamide 4 mg followed by 2 mg after each episode of diarrhea [maximum: 16 mg loperamide daily]) • Dietary modifications (e.g., small lactose-free meals, bananas and rice) • Intake of isotonic fluids 1-1.5 L/day) • Re-assess after 24 hours: - Diarrhea resolving to baseline bowel habits: gradually add solid foods and discontinue or decrease antidiarrheal treatment after 12 h diarrhea-free interval - Diarrhea not resolving: Continue/resume antidiarrheal treatment
Intolerable Grade 2, Grade 2 > 48h, or ≥ Grade 3	• Interrupt study treatment • Ask participant to attend clinic • Rule out infection (e.g., stool sample for culture) - Administer antibiotics as needed, (e.g., if fever or Grade 3-4 neutropenia persists > 24 h) • Administer fluids 1-1.5L/day orally or IV, as appropriate for hydration or to correct electrolyte abnormalities • For Grade 3-4 or complicated lower grade diarrhea consider hospitalization and IV hydration • Re-assess after 24 hour • Diarrhea resolving to baseline bowel habits or Grade ≤ 1: - Consider restarting study treatment at reduced dose • Diarrhea not resolving: - Start and or continue antidiarrheal treatment (e.g., loperamide 4 mg followed by 2 mg after each episode of diarrhea [maximum 16 mg loperamide per day - Consider starting second line antidiarrheal or referral to gastroenterologist

APPENDIX 12 COUNTRY SPECIFIC REQUIREMENTS

COUNTRY SPECIFIC REQUIREMENTS

Country Location	Original language	Country-specific language
Germany, Czech Republic, Romania at all sites where it is mandated by local or country regulations	Section 2 Schedule of Activities, Table 2-1, Screening Procedural Outline, Laboratory Tests	Add "HIV" to the list of laboratory tests
Germany, Czech Republic, Romania at all sites where it is mandated by local or country regulations	Section 6.2 Exclusion Criteria, Exclusion criterion 1 g)	"Known history of testing positive for human immunodeficiency virus (HIV) or known acquired immunodeficiency syndrome (AIDS)"to be replaced with "Positive test for HIV".
Czech Republic	Section 8.1: Discontinuation from Study Treatment	Second paragraph modified to read: "In the case of pregnancy, the investigator must immediately notify the BMS Medical Monitor/designee of this event. In most cases, the study treatment will be permanently discontinued in an appropriate manner (e.g., dose tapering if necessary for participant safety). Please call the BMS Medical Monitor within 24 hours of awareness of the pregnancy." The following statement is deleted: "If the investigator determines a possible favorable benefit/risk ratio that warrants continuation of study treatment, a discussion between the investigator and the BMS Medical Monitor/designee must occur."
Czech Republic, Romania	Section 9.4.1: Clinical Safety Laboratory Assessments	Serology section modified to read: Serum for hepatitis C antibody, HCV RNA, hepatitis B surface antigen (screening only). HIV testing must be performed at sites where mandated by local or country regulations, see Appendix 12)

APPENDIX 13 REVISED PROTOCOL SUMMARY OF CHANGE HISTORY

Overall Rationale for the Revised Protocol 01, 18-Dec-2017

This Global Revised Protocol 01 will stop enrollment into Arm B (nivolumab and ipilimumab combined with cabozantinib triplet regimen) and modify the primary efficacy analysis population to comprise all randomized participants, including International Metastatic RCC Database Consortium (IMDC) favorable risk participants. Recently reported results from the CheckMate 214 (CA209214) study demonstrated that nivolumab in combination with ipilimumab yielded superior overall survival compared to sunitinib in participants with previously untreated, IMDC-intermediate and poor risk mRCC; hence, the nivolumab and ipilimumab combination regimen is anticipated to become a new standard of care in this population. Given these results, it is no longer sufficient for the Arm B triplet regimen to demonstrate superiority over sunitinib, since the nivolumab and ipilimumab combination within the triplet regimen has already been shown to be superior to sunitinib in CA209214.

As the Arm A doublet regimen (nivolumab combined with cabozantinib) does not build upon the nivolumab and ipilimumab combination regimen, there is no *a priori* assumption that the doublet would have superior efficacy to the nivolumab and ipilimumab combination. Given that the nivolumab and ipilimumab combination in CA209214 did not demonstrate a similar clinical benefit in favorable risk participants as was shown in the intermediate and poor risk participants, a key strategic objective for the CA2099ER study is to broaden the first-line mRCC population that may benefit from a nivolumab combination to include all IMDC risk groups, including favorable risk. The Arm C comparator (sunitinib) is indicated in first-line treatment of mRCC across all IMDC risk groups, and it remains an appropriate comparator for the Arm A doublet regimen.

Because the evolving treatment landscape in first-line treatment of mRCC has made it impractical to use the same comparator for the doublet and triplet regimens, the study will cease enrollment to Arm B and will continue as a 2-arm randomization (1:1) comparison between Arm A and Arm C (sunitinib). The primary efficacy analysis population will also be changed to comprise all randomized participants to align with the objective of demonstrating clinical benefit across all IMDC risk groups. The participants at each site who were randomized to Arm B prior to the implementation of this revised protocol will continue the treatment plan and study procedures for Arm B, as specified in the original protocol version 08 Mar 2017.

Secondary items incorporated into this Global Revised Protocol 01 include: (i) to add a Data Monitoring Committee review after 30 participants are treated for 6 weeks, (ii) to adjust, clarify and add exclusion criteria, (iii) to add treatment restrictions, (iv) to clarify criteria associated with hemorrhage with regard to resuming treatment, (v) to specify an additional precaution when sunitinib dosing is resumed, and (vi) to incorporate updated nivolumab clinical program protocol standards.

Tertiary items include: (i) incorporation of Administrative Letter 01and (ii) correction of typographical and grammatical errors.

This Revised Protocol 01 applies in all countries, at all sites, to all future participants enrolled in the study, and where applicable, to all participants currently enrolled in the study

SUMMARY OF KEY CHANGES FOR REVISED PROTOCOL 01		
Section Number & Title	**Description of Change**	**Brief Rationale**
Title Page and Synopsis Protocol Title	In the Protocol Title, Arm B wording is deleted.	Revised Protocol 01 stops randomization to Arm B.
Synopsis and Section 4 Objectives and Endpoints	Arm B wording is deleted. Objectives and endpoints for intermediate/poor risk groups are deleted Exploratory objectives for ipilimumab are deleted	Revised Protocol 01 stops randomization to Arm B. A favorable risk cap of 25% matches the normal frequency in the mRCC population
Synopsis and Section 5.1 Design	Arm B wording is deleted. Number of Arms is revised from 3 to 2, and treatment allocation to 1:1. Sample size is revised Favorable risk cap is adjusted up to 25% Schematic is revised Impact of Revised Protocol 01 on Arm B patients is clarified	Revised Protocol 01 stops randomization to Arm B.
Synopsis, Table 2-3, and Multiple Sections 5.1, 7, 7.1, 7.1.2, 7.4, 8.1.1, 9, 9.5	A paragraph is added to clarify the impact of Revised Protocol 01 on Arm B patients	To clarify the impact of Revised Protocol 01 on Arm B patients in relevant protocol sections.
Procedural Outline Tables 2-2, 2-3 and 2-4	Tumor Tissue row, Notes column - a paragraph is added to clarify procedures associated with Treatment beyond Progression	Procedure description was previously missing
Section 3.1 Study Rationale	Rationale for Arm B is deleted	Revised Protocol 01 stops randomization to Arm B.
Section 3.1.1 Research Hypothesis	Arm B wording is deleted.	Revised Protocol 01 stops randomization to Arm B.
Section 3.2.1.3 Ipilimumab in Renal Cell Carcinoma	Section deleted as no longer relevant	Revised Protocol 01 stops randomization to Arm B.
Section 3.2.1.3 Nivolumab Plus Ipilimumab in Renal Cell Carcinoma	Updated wording from CA209214 is applied. (Section numbering is adjusted from 3.2.1.4 to 3.2.1.3, after deletion of prior section 3.2.1.3)	Revised Protocol 01 stops randomization to Arm B.

SUMMARY OF KEY CHANGES FOR REVISED PROTOCOL 01		
Section Number & Title	**Description of Change**	**Brief Rationale**
Section 3.2.1.4 Cabozantinib Monotherapy in Renal Cell Carcinoma	Paragraph regarding cabozantinib efficacy is updated with ESMO data	The background is updated to reflect the most recent published data
Section 3.3 Benefit Risk Assessment	Arm B wording is deleted. Benefit risk assessment is adjusted to account for the 2 Arm study design	Revised Protocol 01 stops randomization to Arm B.
Section 5.1.1 Data Monitoring Committee	Data Monitoring Committee review after 30 patients are treated for 6 weeks, is added.	Health Authority request (US FDA)
Section 5.2 Number of Participants	Sample size is adjusted	Revised Protocol 01 stops randomization to Arm B. Favorable risk group cap is adjusted to 25%
Section 5.4 Scientific Rationale for Study Design	Arm B wording is deleted. Number of Arms is revised from 3 to 2, and treatment allocation to 1:1. Favorable risk cap is adjusted up to 25%	Revised Protocol 01 stops randomization to Arm B.
Section 5.4.1 Rationale for Open-Label Design	Arm B wording is deleted. Number of Arms is revised from 3 to 2, and treatment allocation to 1:1.	Revised Protocol 01 stops randomization to Arm B.
Section 5.4.2 Rationale for Choice of Primary Endpoint in Intermediate and Poor-Risk Participants	Rationale for adjusting favorable risk cap to 25% is added	Favorable risk group cap is adjusted to 25%
Section 5.5 Justification for Dose	Arm B wording is deleted. Rationale for adjusted cap to favorable risk group, is added	Revised Protocol 01 stops randomization to Arm B. Favorable risk group cap is adjusted to 25%
Section 5.5.2 Dosing for Cabozantinib Doublet and Triplet Therapy in Arm A and Arm B	Arm B wording is deleted and supporting rationale added.	Revised Protocol 01 stops randomization to Arm B.
Section 5.5.4 Rationale for	Arm B wording is deleted.	Revised Protocol 01 stops randomization to Arm B.

SUMMARY OF KEY CHANGES FOR REVISED PROTOCOL 01		
Section Number & Title	**Description of Change**	**Brief Rationale**
Dosing Duration for Nivolumab	With this revised protocol, the option for retreatment on study in subjects who progress after completing 2 years of nivolumab treatment will be removed from the protocol.	Recent evidence suggests that activity of retreatment is very limited.
Synopsis and Section 6.2.1a Exclusion Criteria	Exclusion criteria 6.2.1a revises criteria for CNS metastases that are treated and stable, from at least 3 months to at least 1 month	From 3 months to 1 month after a procedure, the risk of procedure-related complications is low, for patients initiating first line systemic therapy.
Section 6.2.1b Exclusion Criteria	Exclusion criteria 6.2.1b adds 'celiac disease' as an example external trigger.	To address a request from the France Health Authority to add an example of an external trigger.
Section 6.2.1o Exclusion Criteria	Exclusion criteria 6.2.1o revises LMWH criteria for DVT and PE, from at least 6 weeks to at least 3 weeks	To shorten the requirement regarding LMWH treatment
Section 6.2.2e Exclusion Criteria	Exclusion criteria 6.2.2e is revised to major surgery less than 6 weeks, nephrectomy less than 4 weeks, prior to randomization, provided complete wound healing and no ongoing post-operative complications.	To adjust conditions regarding exclusion for major surgery.
Section 6.2.2g Exclusion Criteria	An exclusion is added regarding botanical preparations	To add an exclusion regarding use of botanical preparations
Section 6.2.3f Exclusion Criteria	Exclusion criteria 6.2.3f is clarified	To clarify the requirement for serum creatinine
Section 6.2.3i Exclusion Criteria	Exclusion criteria 6.2.3i adjusts UPCR limit to permit UPCR = 1.5, unless 24-hour urine protein is ≤ 1.5 g	To adjust the UPCR limit.
Section 7.1 Treatment	Clarification of a sentence revising '...approximately 30 minutes...' to read '...at least 30 minutes...'. For consistency, the phrase 'but may be more or less depending on the situation' is deleted.	To clarify a sentence in the Treatment section,
Section 7.4.1.2 Dose Delay Criteria for Cabozantinib	Clarifies that Grade 2 drug-related adverse event or laboratory abnormality (e.g., AST, ALT, total bilirubin) that persists for more than 1 week or worsens despite supportive care management, will delay cabozantinib dosing	To clarify the conditions for delaying cabozantinib dosing, with respect to Grade 2 drug-related AEs and/or lab abnormalities

SUMMARY OF KEY CHANGES FOR REVISED PROTOCOL 01		
Section Number & Title	**Description of Change**	**Brief Rationale**
Section 7.4.1.3 Dose Delay Criteria for Sunitinib	The final paragraph refers the Investigator to the sunitinib label for additional information regarding monitoring of adverse events, in addition to management of adverse events.	To address a request from the France Health Authority regarding the monitoring and management of sunitinib adverse events
Section 7.4.2.2 Dose Reduction and Escalation for Cabozantinib	Participants with asymptomatic Grade 2 drug-related AST, ALT or total bilirubin elevation, or Grade 3 drug-related lipase or amylase abnormalities that are not associated with symptoms or clinical manifestations of pancreatitis may reduce cabozantinib by one dose level, without delaying dosing, at the discretion of the investigator.	To expand the conditions triggering cabozantinib dose reduction
Section 7.4.3.1 Criteria to Resume Nivolumab and Ipilimumab Treatment	Bullet 4 is deleted: Participants with combined Grade 2 AST/ALT AND total bilirubin values meeting discontinuation parameters (Section 8.1.1) should have treatment permanently discontinued.	Revised nivolumab standard.
Section 7.4.3.2 Criteria to Resume Cabozantinib Treatment	Addition of a new bullet; Subjects who develop a pulmonary embolism and/or DVT should have study treatment interrupted until therapeutic anticoagulation is established. Treatment with cabozantinib may be resumed in subjects with pulmonary embolism or DVT if it is determined that the event is uncomplicated and that the subject is deriving clinical benefit from cabozantinib treatment and that anticoagulation does not place them at a significant risk that outweighs the benefit of resuming treatment per discretion of the investigator.	To specify a precaution when cabozantinib dosing is resuming
Section 7.7.1	Addition of a bullet regarding CYP3A4 strong inducers Addition of restriction regarding use of botanical preparations	To add a restrictions regarding use of CYP3A4 strong inducers and botanical preparations
Section 7.7.2.1	Language revises LMWH criteria for DVT and PE, from at least 6 weeks to at least 3 weeks	To shorten the requirement regarding LMWH treatment
Section 8.1.1	Addition of myocarditis to Nivolumab and or Ipilimumab discontinuation criteria	Myocarditis is a newly identified risk

SUMMARY OF KEY CHANGES FOR REVISED PROTOCOL 01		
Section Number & Title	**Description of Change**	**Brief Rationale**
Section 10 Statistical Considerations	Arm B wording is deleted. Number of Arms is revised from 3 to 2, and treatment allocation to 1:1. The analysis is revised from intermediate/poor risk groups to all risk groups Favorable risk cap is adjusted up to 25% The sample size is adjusted from N=1014 to N=580 randomized. Analysis populations, Efficacy Analysis, Safety Analysis and Interim Analysis are revised accordingly.	To stop enrolment into Arm B (Nivolumab, Ipilimumab and Cabozantinib triplet) and to include favorable risk participants (capped at 25%) in the primary data analysis The TKI containing regimen is more appropriately studied in all risk groups.
Appendix 2	Deleted bullets describing records maintained by site	For consistency with revised protocol model document
Appendix 4	Add a barrier method of contraception for patients in Arms A and or B.	For consistency with SmPC label for cabozantinib
Appendix 7	Corrected status definitions of Zubrod-ECOG-WHO scores	For consistency with current descriptions of scores
All	Formatting and typographical corrections	Minor, therefore have not been summarized

STATISTICAL ANALYSIS PLAN

A PHASE 3, RANDOMIZED, OPEN-LABEL STUDY OF NIVOLUMAB COMBINED WITH CABOZANTINIB VERSUS SUNITINIB IN PARTICIPANTS WITH PREVIOUSLY UNTREATED ADVANCED OR METASTATIC RENAL CELL CARCINOMA

PROTOCOL CA209-9ER

VERSION # 1.0

DATE: 29-MAY-2019

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 BACKGROUND AND RATIONALE

Although multiple agents are approved as monotherapies for the treatment of subjects with metastatic renal cell carcinoma (mRCC), the testing of combination therapies, in particular, treatment with immune-checkpoint inhibitors in combination with tyrosine kinase inhibitors (TKIs) has not been fully explored. While single agent therapies have improved outcomes, ongoing drug resistance and disease progression demonstrate an urgent need to find more effective therapies for mRCC subjects.. The CA2099ER trial will include previously untreated subjects with advanced RCC and uses the well-characterized immune checkpoint inhibitor nivolumab in combination with cabozantinib, a known standard-of-care in mRCC subjects. Nivolumab combined with cabozantinib may be an important step forward in evaluating combination regimens which could potentially optimize the management of previously untreated subjects with mRCC.

Cabozantinib was shown to improve PFS and OS compared to everolimus, leading to its regulatory approval in a randomized phase 3 trial in subjects with advanced RCC that had progressed after anti-VEGFR therapy. Subsequently, a randomized phase 2 trial of cabozantinib vs sunitinib has demonstrated an improvement in ORR and PFS in intermediate- and poor-risk subjects with previously untreated mRCC (see Section 3.2.1.4 of the protocol). Cabozantinib has also been demonstrated to have effects on immune cells. In a study of 24 subjects with advanced urothelial carcinoma, cabozantinib treatment resulted in a decrease in circulating Tregs and increased PD-1 expression on Tregs. Low Tregs at baseline were also predictive of improved response to cabozantinib and survival.[1]

Given the promising clinical activity of cabozantinib in previously untreated mRCC and its potential immune effects, combining cabozantinib with nivolumab (in a doublet regimen) is a rational strategy to optimize first-line therapy in mRCC. An ongoing phase 1 study is evaluating both the doublet and triplet (nivolumab and ipilimumab combined with cabozantinib) regimens in subjects with refractory advanced urothelial cancer or other genitourinary tumors, including mRCC, and has defined dosing for both regimens that produces acceptable safety and tolerability (Section 3.2.1.5 of the protocol). CA209-9ER, 2-arm randomized phase 3 trial, will determine if the combination doublet regimen (nivolumab combined with cabozantinib) produces greater clinical benefit than sunitinib, a standard of care agent for subjects with previously untreated mRCC. In addition, this trial will reveal the adverse event profiles, quality of life measures, as well as exploratory biomarkers associated with these different first-line treatment regimens.

Research Hypothesis:

Treatment with nivolumab combined with cabozantinib (doublet regimen) will demonstrate an improvement in PFS per BICR compared to sunitinib monotherapy in subjects with previously untreated mRCC.

Schedule of Analyses:

This study will be monitored by an independent Data Monitoring Committee (DMC). Details are specified in the DMC charter.

The PFS analysis will occur after approximately 9-10 months minimum follow-up on all randomized subjects, which will be triggered by approximately 350 events from Arm A (nivolumab combined with cabozantinib; doublet regimen) and Arm C (sunitinib).

Two interim analyses of OS are planned. The first interim analysis is planned at the time of final PFS analysis and expected after observing 165 deaths among the randomized subjects in Arm A and Arm C (65% of the targeted OS events for final analysis). The second interim analysis is expected after observing 211 deaths among the randomized subjects in Arm A and Arm C (83% of targeted OS events needed for final analysis. The final analysis of OS is expected after observing 211 deaths among the randomized subjects in Arm A and Arm C.

The final PFS analysis will not occur prior to these conditions being met:

- at least 8 months minimum follow-up on all randomized subjects;
- at least 283 PFS events, which provide at least 90% power to detect a HR of 0.68 for PFS of Arm A versus Arm C; and
- at least 149 OS events, which provide 66% power if the observed HR for OS was 0.60.

The expected PFS analysis will occur at approximately 29 months from FPFV. The second interim and final analyses of OS are expected to occur approximately 34 months and 40 months from FPFV.

Secondary endpoints will be analyzed at the time of the final analysis of PFS based on a hierarchical testing strategy. In the event that the interim analysis for superiority of overall survival is positive, final (CSR) analyses will be performed prior to achieving 254 deaths; additional details can be found in section 7.5.7.

2 STUDY DESCRIPTION

Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B (nivolumab + ipilimumab combined with cabozantinib). Subjects previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinical planned events, per protocol.

This is an open label, randomized trial of nivolumab combined with cabozantinib (doublet regimen) versus sunitinib in subjects with previously untreated (first line) advanced or metastatic RCC. Subjects will be randomized between Arm A and Arm C in a 1:1 ratio with approximately 638 subjects (319 per arm) capped at approximate 25% to represent the normal frequency of favorable risk group in mRCC. The rest of the randomized subjects will provide approximately 478 intermediate/poor risk randomized subjects (239 per arm). Subjects will be stratified at the time of randomization by IMDC prognostic score (0 [favorable risk] versus-1-2 [intermediate risk] versus 3-6 [poor risk]), PD-L1 tumor expression (≥ 1% versus < 1% or indeterminate), and region (US/Canada/Western Europe/Northern Europe versus rest of the world [ROW]).

The subject is randomly assigned to 1 of the 2 treatment arms as noted in the study schematic below.

- Arm A (Doublet): Nivolumab 240 mg flat dose IV Q2W + Cabozantinib 40 mg PO QD
  - Nivolumab to be continued until disease progression, unacceptable toxicity, or a maximum treatment of 2 years from the first dose in Cycle 1.
  - Cabozantinib to be continued until disease progression or unacceptable toxicity.

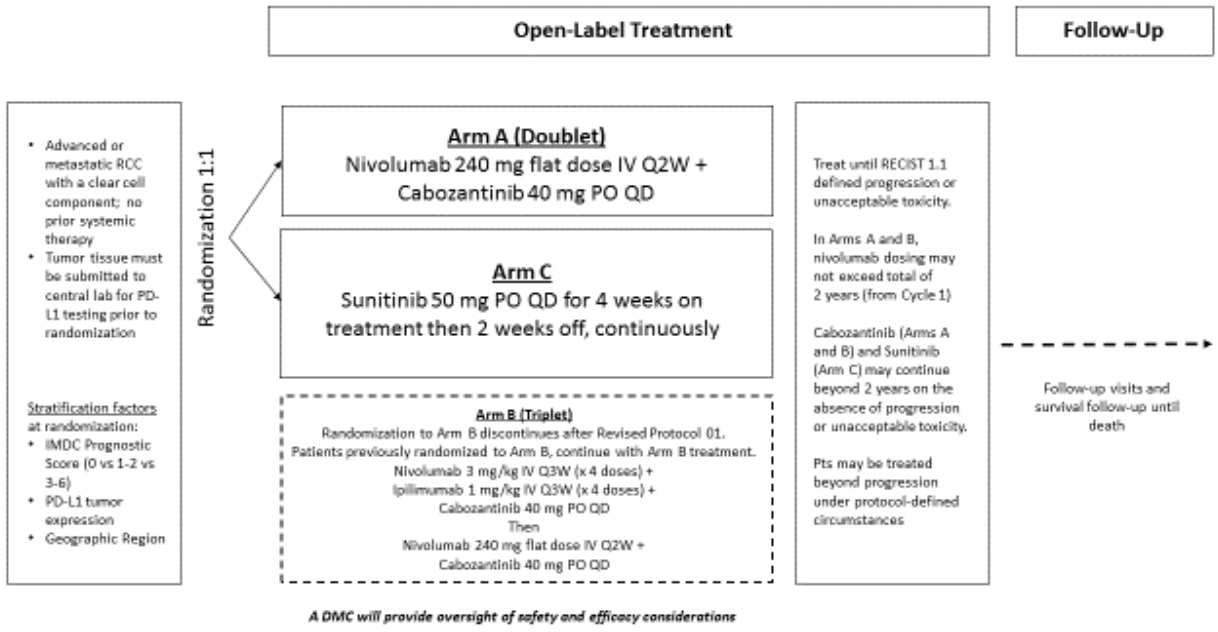
- Arm C: Sunitinib 50 mg PO QD for 4 weeks, followed by 2 weeks off, per cycle. Cycles to be continued until progression or unacceptable toxicity.

- Note - Randomization to Arm B stops with implementation of approved CA2099ER Global Revised Protocol 01. Treatment B continues only for subjects randomized to Arm B prior to implementation of Global Revised Protocol 01.
  - Arm B (Triplet): Nivolumab 3mg/kg IV + Ipilimumab 1 mg/kg IV, both Q3W x 4 doses + Cabozantinib 40 mg PO QD
    - Then Nivolumab 240 mg flat dose IV Q2W + Cabozantinib 40 mg PO QD.
    - Nivolumab to be continued until disease progression, unacceptable toxicity, or a maximum of 2 years from the first dose in Cycle 1.
    - Cabozantinib to be continued until progression or unacceptable toxicity.

Once randomized subjects in Arm A will continue nivolumab until progression, unacceptable toxicity, withdrawal of consent, or a maximum of 2 years from the first dose in Cycle 1, whichever occurs first. Cabozantinib (Arm A) may be continued until progression, unacceptable toxicity, or withdrawal of consent, whichever occurs first, and may extend beyond 2 years from the first dose in Cycle 1.

The study design schematic is presented in the figure below.

Figure 2-1: Study Design Schematic



2.1 Treatment Assignment

After the subject's initial eligibility is established and informed consent has been obtained, the subject must be enrolled into the study by accessing an Interactive Response Technologies web-based system (IRT) to obtain the subject number. All subjects will be centrally randomized using an Interactive Response Technology (IRT). Before the study is initiated, each user will receive log in information and directions on how to access the IRT.

Every subject that signs the informed consent form must be assigned a subject number in IRT. The investigator or designee will register the subject for enrollment by following the enrollment procedures established by BMS. The following information is required for enrollment:

- Date that informed consent was obtained
- Date of birth
- Sex at birth.

Once enrolled in IRT, enrolled subjects that have met all eligibility criteria will be ready to be randomized through the IRT. The following information is required for subject randomization:

- Subject number
- Date of birth
- IMDC Prognostic Score (0 versus 1-2 versus 3-6)
- Region (US/Canada/W Europe/N Europe versus ROW)
- PD-L1 tumor expression (≥ 1% versus < 1% or indeterminate)

Subjects meeting all eligibility criteria will be randomized in a 1:1 ratio to Arm A (nivolumab combined with cabozantinib) or Arm C (sunitinib), stratified by the following factors:

- IMDC Prognostic Score: 0 versus 1-2 versus 3-6
- Region: US/Canada/W Europe/N Europe versus ROW
- PD-L1 tumor expression: ≥ 1% versus < 1% or indeterminate

The randomization procedures will be carried out via permuted blocks within each stratum.

2.2 Blinding and Unblinding

This is an open label study.

2.3 Protocol Amendments

Table 2.3-1: Protocol Amendments

Document	Date of Issue	Summary of Change
Revised Protocol 02	06-MAY-2019	Major Changes: • Revised protocol 02 adjusts the timing of the PFS and OS interim analyses with modified hypothesized OS hazard ratio (HR). The number of randomized subjects is increased.

Table 2.3-1: Protocol Amendments

Document	Date of Issue	Summary of Change
Revised Protocol 01	18-DEC-2017	• The interim analysis for ORR is removed, resulting in revised overall alpha for PFS and OS endpoints. • No change in eligibility or study procedure. • Clinical data for nivolumuab + ipilimumab in RCC has been updated. Other changes include more detail on PRO measures and updates to align with BMS standards for the nivolumab program. Primary revisions: (i) To stop enrollment into Arm B (nivolumab, ipilimumab and cabozantinib triplet) and (ii) to include favorable risk subjects (capped at 25%) in the primary data analysis. Secondary items include: (i) to add a Data Monitoring Committee review after 30 subjects are treated for 6 weeks, (ii) to adjust, clarify and add exclusion criteria, (iii) to add treatment restrictions, (iv) to clarify criteria associated with hemorrhage with regard to resuming treatment, (v) to specify an additional precaution when sunitinib dosing is resumed, and (vi) to apply newly updated Sponsor standards for nivolumab clinical protocols. Tertiary items include (i) incorporation of Administrative Letter 01 and (ii) correction of typographical and grammatical errors.

2.4 Data Monitoring Committee

An independent Data Monitoring Committee (DMC) has been established to provide oversight of safety and efficacy considerations, study conduct, and risk-benefit ratio. Following review, the DMC will recommend continuation, modification, or discontinuation of this study based on reported safety and efficacy data. Details of DMC responsibilities and procedures are specified in the DMC charter. Representatives of the Sponsor will serve only as coordinators of the committee, without having full member responsibilities or privileges. In addition, the Sponsor will independently review safety data in a blinded manner during the conduct of this trial to ensure that any safety issues are identified and addressed.

The DMC will conduct the first review of the safety data after at least 30 subjects are treated and followed for at least 6 weeks. The DMC will conduct its second review of the safety data after at least 75 subjects are treated and followed for at least 6 weeks. The DMC will conduct its review of the safety data focusing on the initial approximately 12 Japanese subjects (6 per arm) treated and followed for at least 4 weeks. The DMC will then review safety and the available efficacy data pertaining to primary endpoint to evaluate safety in the context of benefit, every six months thereafter.

The DMC will also review the formal final analysis of PFS (as per BICR) and first interim analysis of superiority of OS scheduled at around 29 months from FPFV. A second interim analysis of OS

will be at around 34 months from FPFV. BMS will remain blinded to OS interim analyses unless DMC decides to disclose the formal interim analysis to BMS.

Details of the interim analyses can be found in section 7.5.7.

2.5 Blinded Independet Central Review

A blinded independent central review (BICR) committee has been established to provide an independent imaging review of images obtained in subjects participating in this study. Details of BICR responsibilities and processes may be found in the BICR Charter. The BICR determined PFS and ORR endpoints will be utilized as a part of primary and secondary efficacy analyses.

3 OBJECTIVES

3.1 Primary

- To compare PFS per BICR of nivolumab combined with cabozantinib (Arm A: doublet) with sunitinib (Arm C) in all randomized subjects.

3.2 Secondary

- To compare overall survival (OS) of Arm A with Arm C in all randomized subjects.
- To evaluate the objective response rate (ORR) of Arm A with Arm C per BICR in all randomized subjects.
- To assess overall safety and tolerability in all treated subjects.

3.3 Exporatory

- To explore potential predictive biomarkers of clinical response to nivolumab and cabozantinib combination.
- To evaluate health related quality of life (HRQoL).
- To characterize the pharmacokinetics of nivolumab and cabozantinib and explore exposure response relationships, if applicable.
- To characterize the immunogenicity of nivolumab.
- To assess PFS after next line of treatment (PFS2) in each arm.

4 ENDPOINTS

4.1 Primary Endpoints

Progression-free survival (PFS) is the primary endpoint. Two definitions are used for the analysis of PFS. The primary definition accounts for subsequent therapy by censoring at the last evaluable tumor assessment on or prior to the date of subsequent therapy. The secondary definition is irrespective of subsequent therapy and does not account for subsequent therapy.

Clinical deterioration in the absence of unequivocal evidence of progression (per RECIST v1.1 criteria) is not considered progression for purposes of determining PFS.

PFS rate at time *T* is defined as the probability that a subject has not progressed and is alive at time *T* following randomization. PFS rates at fixed time points (e.g. 6 months, depending on the minimum follow-up) are defined as the probability that a subject has not progressed and is alive at time *T* following randomization.

The first on-study tumor assessment is scheduled to be conducted at 12 weeks (± 1 week) following randomization. Subsequent tumor assessments are scheduled every 6 weeks (± 1 week) up to week 60, then every 12 weeks (± 2 weeks) until disease progression.

4.1.1 Primary Definition of Progression-Free Survival (Accounting for Subsequent Therapy)

The primary definition of PFS (PFS truncated at subsequent therapy, which includes anti-cancer therapy, tumor directed radiotherapy, or tumor directed surgery) is defined as the time between the date of randomization and the date of first documented tumor progression, based on BICR assessments (per RECIST v1.1 criteria), or death due to any cause, whichever occurs first.

Subjects who die without a reported progression will be considered to have progressed on the date of their death. The following censoring rules will be applied for the primary definition of PFS:

- Subjects who did not progress or die will be censored on the date of their last evaluable tumor assessment.
- Subjects who did not have any on study tumor assessments and did not die will be censored on their date of randomization.
- Subjects who receive subsequent anti-cancer therapy prior to documented progression will be censored at the date of the last evaluable tumor assessment conducted on or prior to the date of initiation of the subsequent anti-cancer therapy.
- Subjects who did not have a documented progression and received subsequent anti-cancer therapy will be censored at the date of the last evaluable tumor assessment conducted on or prior to the initiation of the subsequent anti-cancer therapy.

Censoring rules for the primary definition of PFS (PFS truncated at subsequent therapy) are presented as follows and in Table 4.1.1-1.

Figure 4.1.1-1: PFS Primary Definition

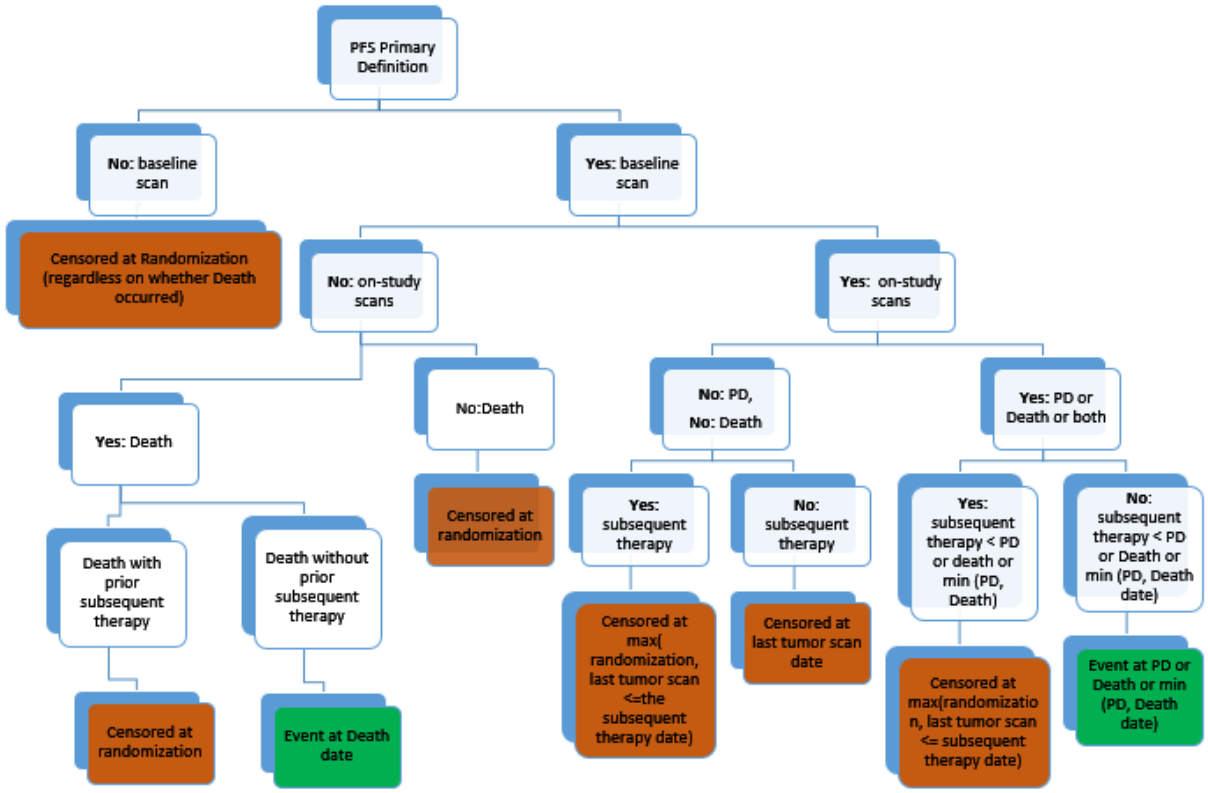


Table 4.1.1-1: Censoring Scheme Used in Primary Definition of PFS

Situation	Date of Progression or Censoring	Outcome
No baseline tumor assessments*	Date of randomization	Censored
No on study tumor assessments and no death*	Date of randomization	Censored
Subsequent anti-cancer therapy started without death or progression per RECIST v1.1 reported prior or on the same day	Date of last evaluable tumor assessment prior to or on the date of initiation of the subsequent anti-cancer therapy	Censored
Documented progression per RECIST v1.1 and no new anti-cancer started before	Date of the first documented progression per RECIST v1.1 (excludes clinical progression)	Progressed
No progression and no death, and no new anti-cancer therapy started	Date of last evaluable tumor assessment	Censored
Death without progression per RECIST v1.1 and no new anti-cancer started before	Date of death	Progressed

* Tumor assessments and death if any, occurring after start of subsequent anti-cancer therapy are not considered.

4.1.2 Secondary Definition of Progression Free Survival (Irrespective of Subsequent Therapy)

The secondary definition of PFS (ITT definition) is defined as the time between the date of randomization and the date of first documented tumor progression, based on BICR assessments (per RECIST v1.1 criteria), or death due to any cause, whichever occurs first.

Subjects who die without a reported progression will be considered to have progressed on the date of their death. The following censoring rules will be applied for the secondary definition of PFS:

- Subjects who did not progress or die will be censored on the date of their last evaluable tumor assessment.
- Subjects who did not have any on study tumor assessments and did not die will be censored on their date of randomization.

Censoring rules for the secondary definition of PFS (ITT definition) are presented as follows and in Table 4.1.2-1.

Figure 4.1.2-1: PFS Secondary Definition

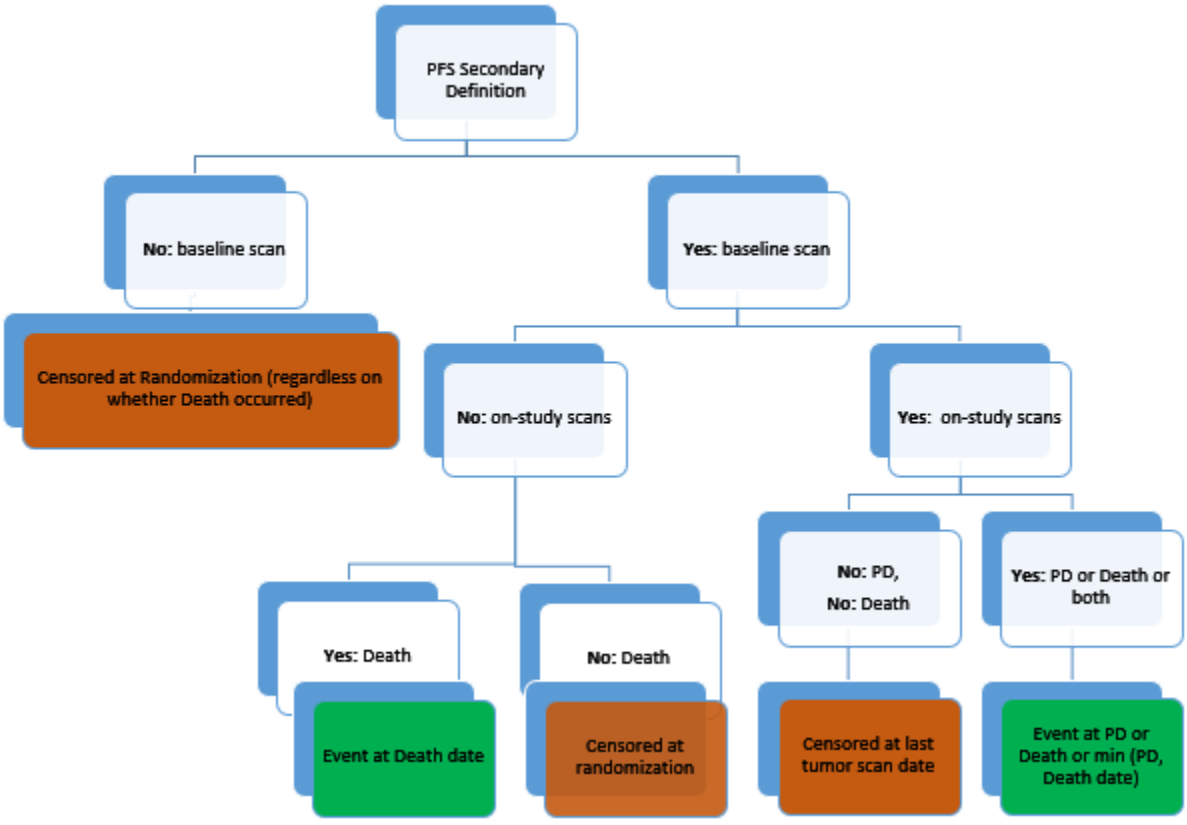


Table 4.1.2-1: Censoring Scheme for Secondary Definition of PFS

Situation	Date of Progression of Censoring	Outcome
No baseline tumor assessment	Date of randomization	Censored
No on-study tumor assessments and no death	Date of randomization	Censored
Documented progression per RECIST v1.1	Date of first documented progression per RECIST v1.1 criteria (excludes clinical progression)	Progressed
No progression and no death	Date of last evaluable tumor assessment	Censored
Death without progression per RECIST v1.1	Date of death	Progressed

4.2 Secondary Endpoints

4.2.1 Overall Survival

Overall survival (OS) is defined as the time from randomization to the date of death from any cause. For subjects that are alive, their survival time will be censored at the date of last contact date (or "last known alive date"). Overall survival will be censored at the date of randomization for subjects who were randomized but had no follow-up.

Follow-up visit #1 (FU1) should occur 30 days from the last dose and follow-up visit #2 (FU2) occurs approximately 100 days from last dose of study drug. After FU2, survival follow-up will be conducted every 3 months.

4.2.2 Objective Response Rate

Objective Response Rate (ORR) is defined as the number of randomized subjects who achieve a best response of confirmed complete response (CR) or confirmed partial response (PR) based on BICR assessments (using RECIST v1.1 criteria) divided by the number of all randomized subjects. Best Overall Response (BOR) is defined as the best response, as determined by the BICR, recorded between the date of randomization and the date of objectively documented progression per RECIST v1.1 criteria or the date of subsequent therapy (including tumor-directed radiotherapy and tumor-directed surgery), whichever occurs first. For subjects without documented progression or subsequent therapy, all available response designations will contribute to the BOR determination. Confirmation of response is required at least 4 weeks after the initial response.

4.2.2.1 Time to Response

Time to Response (TTR) is defined as the time from randomization to the date of the first confirmed documented response (CR or PR), as assessed by the BICR. TTR will be evaluated for responders (confirmed CR or PR) only.

4.2.2.2 Duration of Response

Duration of Response (DOR) is defined as the time between the date of first confirmed documented response (CR or PR) to the date of the first documented tumor progression as determined by the BICR (per RECIST v1.1 criteria), or death due to any cause, whichever occurs first. Subjects who start subsequent therapy without a prior reported progression will be censored at the last evaluable tumor assessments prior to initiation of the subsequent anti-cancer therapy. Subjects who die without a reported prior progression will be considered to have progressed on the date of their death. Subjects who neither progress nor die, DOR will be censored on the date of their last evaluable tumor assessment. DOR will be evaluated for responders (confirmed CR or PR) only.

4.3 Safety Endpoints

The assessment of safety will be based on the incidence of adverse events (AEs), serious adverse events (SAEs), adverse events leading to discontinuation, adverse events leading to dose modification, select adverse events (select AEs) for EU/ROW Submissions, immune-mediated AEs (IMAEs) for US Submission, other events of special interest (OEOSI), and deaths. The use of immune modulating concomitant medication will be also summarized. In addition clinical laboratory tests, and immunogenicity (i.e. development of anti-drug antibody) will be analyzed.

4.4 Exploratory Endpoints

4.4.1 Biomarkers

Biomarkers potentially associated with clinical endpoints will be measured by analyzing tumor and blood samples. Biomarker endpoints include, but are not limited to, single-nucleotide polymorphisms (SNPs), proteins in tumor specimens and serum, and immune cell populations.

4.4.1.1 PD-L1 Expression

PD-L1 expression is defined as the percent of tumor cells membrane staining in a minimum of 100 evaluable tumor cells per validated Dako PD-L1 immunohistochemistry (IHC) assay. This is referred to as quantifiable PD-L1 expression. If the PD-L1 staining could not be quantified, it is further classified as:

1) Indeterminate: Tumor cell membrane staining hampered for reasons attributed to the biology of the tumor tissue sample and not because of improper sample preparation or handling.
2) Not evaluable: Tumor tissue sample was not optimally collected or prepared and PD-L1 expression is neither quantifiable nor indeterminate. Not evaluable can be determined from H&E process before the tumor biopsy specimen is sent for PD-L1 evaluation or from the H&E process during PD-L1 evaluation.

Subjects with missing PD-L1 expression are subjects with no tumor tissue sample available for evaluation.

PD-L1 expression will be collected in the IRT as well as in the clinical database. Statistical analysis using PD-L1 expression will be solely based on PD-L1 expression data from clinical database.

Efficacy endpoints defined above (PFS and ORR by BICR, OS) will be analyzed by PD-L1 expression to explore the association of PD-L1 expression on tumor and/or tumor associated immune cells (TAIC) with clinical benefit.

4.4.2 Clinical Outcomes Assessments

The FSKI-19 and EQ-5D-3L patient-reported outcomes will be collected and analyzed in this study.

4.4.2.1 FKSI-19

The NCCN FKSI-19 is a 19-item scale that measures tumor specific HrQoL in kidney cancer subjects. The FKSI-19 uses 5 Likert-type response categories that range from "not at all" to "very much." Subjects are asked to circle the response category that best characterizes their response over the last 7 days on 19 items that include symptoms such as lack of energy, fatigue, appetite, coughing, shortness of breath, pain, nausea, and ability to work.The instrument yields a total score and three subscale scores: Disease Related Symptoms (DRS), Treatment Side Effects (TSE), and Functional Well Being (FWB). A higher score indicates fewer symptoms.

If there are missing items, subscale scores can be prorated according to the standard FACT (Functional Assessment of Cancer Therapy) scoring methodology. This can be done using the following formula:

$$\boldsymbol{Prorated\ Subscale\ Score} = \frac{[\boldsymbol{Sum\ of\ Item\ Scores}]\ \boldsymbol{x}\ [\boldsymbol{N\ of\ Items\ in\ Subscale}]}{[\boldsymbol{N\ of\ Items\ Answered}]}$$

When there are missing data, prorating by subscale in this way is acceptable as long as more than 50% of the items were answered. The total score is then calculated as the sum of the un-weighted subscale scores.

Table 4.4.2.1-1: Time Windows for FKSI-19 and EQ-5D-3L Assessments

Nominal Time-Point (Cycle)	Time Window
Nivolumab+Cabozantinib (Arm A) treatment group	
Baseline (Cycle 1)	Prior to first dose on Day 1
Week 2 (Cycle 2)	Nominal Day 15 (Day 2 thru Day 24, inclusive of Day 15 + 9 days)
Every 2 weeks thereafter (Cycles 3+)	Nominal Days 29+ (+9 days/-5 days, inclusive)

Table 4.4.2.1-1: Time Windows for FKSI-19 and EQ-5D-3L Assessments

Nominal Time-Point (Cycle)	Time Window
Nivolumab+Ipilimumab+Cabozantinib (Arm B) treatment group	
Baseline (Cycle 1)	Prior to first dose on Day 1
Week 3 (Cycle 2)	Nominal Day 22 (Day 2 thru Day 38, inclusive of Day 22 + 16 days)
Every 3 weeks up to Week 9 (Cycles 3-4)	Nominal Days 43 and 64 (+16 days/-5 days, inclusive)
Week 12 (Cycle 5)	Nominal Day 85 (+9 days/-5 days, inclusive)
Every 2 weeks thereafter (Cycles 6+)	Nominal Days 99+ (+9 days/-5 days, inclusive)
Sunitinib (Arm C) treatment group	
Baseline (Cycle 1)	Prior to first dose on Day 1
Week 6 (Cycle 2)	Nominal Day 43 (Day 2 thru Day 80, inclusive of Day 43 + 37 days)
Every 6 weeks thereafter (Cycles 3+)	Nominal Days 85+ (+9 days/-5 days, inclusive)
For all treatment groups	
Follow Up:	
Follow-Up 1	30 days from the last dose (± 7 days) or coincide with the date of discontinuation (± 7 days) if date of discontinuation is greater than 42 days after last dose.
Follow-Up 2	100 days (± 7 days) from last dose of study treatment. Subjects must be followed for at least 100 days after last dose of study treatment.
Survival Follow-Up i [a] (i = 1, 2, 3, 4, ...)	3 months (± 14 Days) after Follow-up Visit 2, and subsequent survival follow-up visits every 3 months (± 14 days).

[a] Survival Follow-Up is only relevant for EQ-5D, not FKSI-19 assesments.

4.4.2.2 EuroQoL EQ-5D-3L

Subjects' reports of general health status will be assessed using the EuroQoL Group's EQ-5D-3L. EQ-5D-3L essentially has 2 components: the descriptive system and the visual analogue scale (VAS).

The instrument's descriptive system consists of 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has 3 levels, reflecting "no health problems," "moderate health problems," and "extreme health problems." A dimension for which there are no problems is said to be at level 1, while a dimension for which there are extreme problems is said to be at level 3. Thus, the vectors 11111 and 33333 represent the best health state and the worst health state, respectively, described by the EQ-5D-3L. Altogether, the instrument describes $3^5 = 243$ health states. Empirically derived weights can be applied to an individual's

responses to the EQ-5D-3L descriptive system to generate an index measuring the value to society of his or her current health. Such preference-weighting systems have been developed for the UK, US, Spain, Germany, and numerous other populations. For this study, EQ-5D-3L utility index values will be computed using a scoring algorithm based on the United Kingdom Time-Trade-Off (UK TTO) value set[2]

In addition, the EQ-5D-3L includes a VAS, which allows respondents to rate their own current health on a 101-point scale ranging from 0=”worst imaginable” health to 100=“best imaginable” health state [3].

A change from baseline of 0.08 for the EQ-5D-3L utility index score and of 7 for the EQ-5D-3L VAS are considered minimally important differences for the EQ-5D-3L[4].

All questionnaires completed at baseline and on-study will be assigned to a time-point according to the windowing criteria in Table 4.4.2.1-1 and included in the analysis. In case a subject has two on-study assessments within the same window, the assessment closest to the time-point will be used. And, in the case of two assessments at a similar distance to the time-point, the latest one will be chosen. In the event where the subject has no assessment at all in a specific window, the observation will be treated as missing for that time-point.

4.4.3 Pharmacokinetics

PK will be measured by the serum concentration of nivolumab and/or ipilimumab and/or cabozantinib. Samples will be collected to characterize pharmacokinetics of nivolumab and to explore exposure-safety and exposure-efficacy relationships.

4.4.4 Immunogenicity

Serum samples collected will be analyzed by a validated immunogenicity assay. Selected serum samples may be analyzed by an exploratory orthogonal method that measures anti-nivolumab.

In addition, ad hoc serum samples designated for pharmacokinetic or biomarker assessments may also be used for immunogenicity analysis if required (e.g., insufficient volume for complete immunogenicity assessment or to follow up on suspected immunogenicity related AE).

Further details on immunogenicity background and rationale, definitions, population for analyses and endpoints are described in APPENDIX 3.

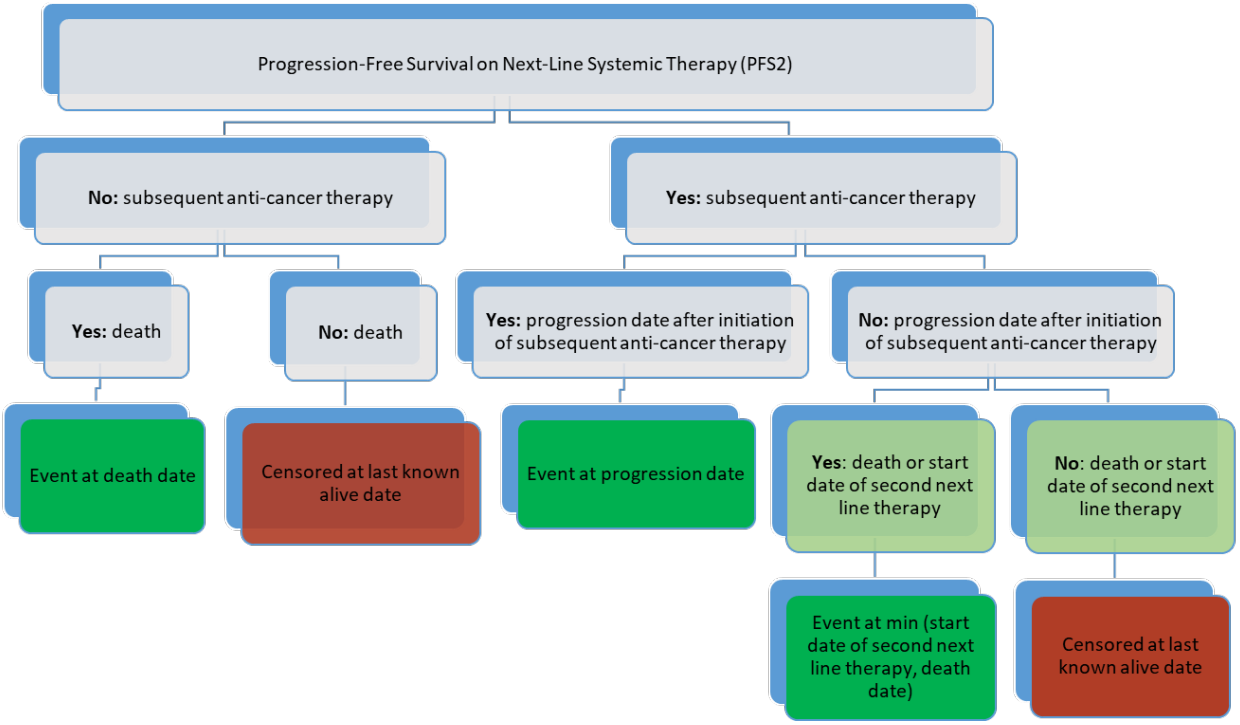
4.4.5 PFS2

PFS on next-line therapy (PFS2) is defined as the time from randomization to objectively documented progression, per investigator assessment, after the next line of therapy or to death from any cause, whichever occurs first. Subjects who were alive and without progression after the next line of therapy will be censored at last known alive date.

The following censoring rules will be applied for PFS2:

- Subjects who did not receive subsequent anti-cancer therapy (i.e. second-line therapy):
  - Subjects who died, the death date is the event date;
  - Else the subject's PFS2 is censored at the last known alive date.
- Subjects who received subsequent anti-cancer therapy (i.e. second-line therapy):
  - Subjects who had a disease progression after the start of subsequent anti-cancer therapy, this disease progression date is the event date;
  - Else if a subject died or start of second next line therapy, the date of min (death, start date of second next line therapy) is the event date;
  - Else the subject's PFS2 is censored at the last known alive date.

Figure 4.4.5-1: PFS2 Definition



5 SAMPLE SIZE AND POWER

The sample size calculations of this study summarized below are based on the target randomized subjects in Arm A and Arm C only. That is, the total randomized subjects will be higher than the target randomized subjects due to randomized subjects into Arm B prior to implementation of CA2099ER Global Revised Protocol 01, which stopped further randomization into Arm B.

The sample size of this study accounts for the primary endpoint of progression-free survival (PFS) per BICR in all randomized subjects. Assuming a 25% screen failure rate, it is expected that approximately 850 subjects will need to be enrolled in order to randomize 638 subjects (319 per arm) in a 1:1 ratio. To represent the normal frequency of the favorable risk group in mRCC, the favorable risk subjects are capped at approximate 25%; thus, at most 212 favorable risk subjects (106 per arm) will be enrolled to randomize 160 favorable risk subjects in a 1:1 ratio. The rest of the enrolled participants will provide approximately 478 intermediate/poor risk randomized subjects (239 per each arm).

The overall alpha for this study is 0.05 (two-sided). PFS will be evaluated for treatment effect at an alpha of 0.05 (two-sided), with at least 95% power. No interim analysis of PFS is planned. OS will be evaluated for treatment effect at an alpha level of 0.05 (two-sided) with 80% power, accounting for two formal interim analyses to assess efficacy.

Sample Size Justification for Primary PFS Endpoint

The primary endpoint of PFS per BICR of Arm A versus Arm C analysis conducted on all randomized subjects. The PFS analysis will occur after approximately 9-10 months minimum follow-up on all randomized subjects, which will be triggered by approximately 350 events from Arm A and Arm C. The 350 PFS events provide at least 95% power to detect a HR of 0.68 for PFS of Arm A versus Arm C with a type I error of 0.05 (two-sided). The HR of 0.68 corresponds to a 47% increase in the median PFS, assuming a median PFS of 18.2 months for Arm A and 12.4 months for Arm C. It is projected that an observed HR of 0.811 or less, which corresponds to a 2.89 month or greater improvement in median PFS (12.4 versus 15.3 months), would result in a statistically significant improvement in PFS for the Arm A versus Arm C comparison.

If the formal analysis of PFS among all randomized subjects is statistically significant, the formal interim analysis of OS among all randomized subjects will be tested, as per hierarchical testing procedure.

Sample Size Computation for Secondary OS Endpoint

The secondary endpoint of OS in all randomized subjects specifies the comparison of Arm A versus Arm C. Among all randomized subjects, approximately 254 events (ie, deaths) in Arm A and Arm C provides at least 80% power to detect a HR of 0.70 for OS of Arm A and Arm C with an overall type 1 error of 0.05 (two-sided) for each test. The HR of 0.70 corresponds to a 43% increase in the median OS, assuming a median OS of 47.1 months for Arm A and 33 months for Arm C.

Two formal interim analyses of OS are planned for this study. The first interim analysis is planned at the time of final PFS analysis and it is expected to observe 165 OS events (65% of the targeted OS events for final analysis) and the second interim analysis is planned to occur after observing approximately 211 events (83% of targeted OS events needed for final analysis). The stopping boundaries at interim and final analyses will be derived based on the number of deaths using O'Brien and Fleming α spending function. For example, with 165, 211, and 254 observed events in Arm A and Arm C at the first interim, second interim, and final analyses, the respective stopping

boundaries would be α=0.011 (two-sided), α=0.025 (two-sided), and α=0.041 (two-sided). If the first interim analysis is performed exactly at 165 deaths, it is projected that an observed HR of 0.673 or less, which corresponds to a 16.0 month or greater improvement in median OS (33 versus 49 months), would result in a statistically significant improvement in OS for the Arm A versus Arm C comparison. At the second interim analysis with 211 deaths, it is projected that an observed HR of 0.734 or less, which corresponds to a 12.0 month or greater improvement in median OS (33 versus 45 months), would result in a statistically significant improvement in OS for the Arm A versus Arm C comparison. At the time of final OS analysis when there are 254 deaths, it is projected that an observed HR of 0.774 or less, which corresponds to a 9.6 month or greater improvement in median OS (33 versus 42.6 months), would result in a statistically significant improvement in OS for the Arm A versus Arm C comparison.

Assuming a constant accrual rate (an average rate of 3 subjects/month in the first 4 months, afterwards an average rate of 42 subjects/month), the accrual will take approximately 19 months. The final PFS analysis will not occur prior to these conditions being met:

- at least 8 months minimum follow-up on all randomized subjects;
- at least 283 PFS events, which provide at least 90% power to detect a HR of 0.68 for PFS of Arm A versus Arm C; and
- at least 149 OS events, which provide 66% power if the observed HR for OS was 0.60. (Note that if the analysis of first interim analysis OS takes place with 149 OS events, the alpha spending for the OS comparison would be 0.007 with a critical HR=0.643.)

This expected PFS analysis will occur at approximately 29 months from FPFV. The second interim and final analyses of OS are expected to occur approximately 34 months and 40 months from FPFV, respectively. Table 5-1 summarizes the results of these calculations.

Table 5-1: Summary of Sample Size Parameters and Schedule of Analyses

Primary/Secondary Endpoints	PFS (Primary)	OS (Secondary)
Primary analysis population	All Randomized Subjects	
Accrual rate per month for all randomized population	3 subjects/month in the first 4 months, afterwards an average rate of 42 subjects/month	
Power	95%	80%
Alpha	0.05 2-sided	0.05 2-sided (0.011 at IA1, 0.025 at IA2, 0.041 at FA)
Hypothesized median control vs exp (months)	12.4 vs 18.2	33 vs 47.1
Hypothesized hazard ratio	0.68	0.70

Table 5-1: Summary of Sample Size Parameters and Schedule of Analyses

Primary/Secondary Endpoints	PFS (Primary)	OS (Secondary)
Critical hazard ratio (observed hazard ratio at which a statistically significant difference would be observed) / Difference in median (months) Corresponding to a minimal clinically significant effect size (FA)	0.811 / 2.89	0.774 / 9.6
Critical HR at interim analysis-1(IA1) /effect size	N/A	0.673 / 16.0
Expected number of event for IA1 (percentage of target events)	N/A	165 (65%)
Timing of IA1 from FPFV (months)	N/A	29
Critical HR at interim analysis-2(IA2) /effect size	N/A	0.734 / 12.0
Expected number of event for IA2 (percentage of target events)	N/A	211 (83%)
Timing of IA2 from FPFV (months)	N/A	34
Accrual duration (months)	19	19
Timing of final analysis (FA) from FPFV (months)	29	40
Sample size	638	638
Target number of events (Event Goal)	350	254

6 STUDY PERIODS, TREATMENT REGIMENS AND POPULATIONS FOR ANALYSES

6.1 Study Periods

- Baseline period:
  - Baseline evaluations or events will be defined as evaluations or events that occur before the date and time of the first dose of study treatment. Evaluations (laboratory tests, pulse oximetry and vital signs) on the same date and time of the first dose of study treatment will be considered as baseline evaluations. Events (AEs) on the same date and time of the first dose of study treatment will not be considered as pre-treatment events.

- In cases where the time (onset time of event or evaluation time and dosing time) is missing or not collected, the following definitions will apply:
  - Pre-treatment AEs will be defined as AEs with an onset date prior to but not including the day of the first dose of study treatment;
  - Baseline evaluations (laboratory tests, pulse oximetry and vital signs) will be defined as evaluations with a date on or prior to the day of first dose of study treatment.
- If there are multiple valid assessments on or prior to the first dose of study treatment:
  - For laboratory tests, the latest non missing labs value on or before first dose date (and time if collected) will be used as the baseline in the analyses. For 'LIPASE' and 'GLUCOSE', for treated subjects only, the last predose assessment with non-missing toxicity grade will be considered as baseline. If multiple assessments exist with the same collection date (and time if collected) and entry date and time, then the first observation is used as baseline.
  - For PD-L1, among the records prior to or on first dose date (and time if collected), identify first those with quantifiable test result. If there are no records with quantifiable test result, then select those with indeterminant result ("INDETERMINATE"). If there are no records with indeterminant test result, then select those with unavailable result ("NOT EVALUABLE). If there are no records with unavailable test result, then select those with not reported or not available result (all other records). The latest record will be used as the baseline in the analyses. If there is more than one record for the latest date, then choose the one with the greatest specimen ID.
  - For Anti-Drug Antibody (ADA), the record related to the most recent assessment among those records where date (and time if collected) of Nivolumab immunoglobulin (IMG) assessment is less than or equal to the date (and time if collected) of the first Nivolumab dose date.

- Post baseline period:
  - On-treatment AEs will be defined as AEs with an onset date and time on or after the date and time of the first dose of study treatment (or with an onset date on or after the day of first dose of study treatment if time is not collected or is missing). For subjects who are off study treatment, AEs will be included if event occurred within a safety window of 30 days (or 100 days depending on the analysis) after the last dose of study treatment. No "subtracting rule" will be applied when an AE occurs both pre-treatment and post-treatment with the same preferred term and grade.
  - On-treatment evaluations (laboratory tests, pulse oximetry and vital signs) will be defined as evaluations taken after the day (and time, if collected and not missing) of first dose of study treatment. For subjects who are off study treatment, evaluations should be within a safety window of 30 days (or 100 days depending on the analysis) after the last dose of study treatment.

- Late-emergent drug-related AEs will be defined as drug-related AEs with an onset date greater than 100 days after the last dose of study treatment in subjects who are off study treatment.

6.2 Treatment Regimens

The treatment group "**as randomized"** will be retrieved from the IRT system:

- Arm A: Experimental arm: nivolumab + cabozantinib
- Arm B: Experimental arm: nivolumab + ipilimumab + cabozantinib
- Arm C: Control arm: sunitinib

The treatment group "**as treated**" will be the same as the arm as randomized by IRT. However, if a subject received the incorrect drug for **the entire period** of treatment, the subject's treatment group will be defined as the incorrect drug the subject actually received.

Unless otherwise specified, the safety analysis will be based on the treatment group "as treated".

Unless otherwise specified, the efficacy analysis will be based on the treatment group "as randomized".

6.3 Populations for Analyses

All analyses will be performed using the treatment arm as randomized (intent to treat), with the exception of dosing and safety, for which the treatment arm as received will be used. For purposes of analysis, the following populations are defined in Table 6.3-1, and all populations for analyses given in this table refer to those subjects in Arm A and Arm C. Those subjects who randomized to Arm B prior to Revised Protocol 01 will be considered as part of the population of interest for descriptive summary of efficacy and safety analyses.

Table 6.3-1: Populations for Analyses

Population	Description
All Enrolled Subjects	All subjects who sign informed consent and were registered into the IRT.
All Randomized Subjects	All subjects who were randomized will be used for analyses of demography, protocol deviations, baseline characteristics, primary efficacy analysis, secondary efficacy analyses, and outcome research analysis which will be performed for this population.
All Treated Subjects	All subjects who received at least one dose of any study medication. This is the primary population for exposure and safety analyses.
Intermediate/Poor Risk Subjects	All subjects who were randomized with baseline IMDC prognostic score ≥1 at the time of randomization (per IRT). This population will be used for subset analyses of demography, protocol deviations, baseline characteristics, primary efficacy analysis, and secondary efficacy analyses on intermediate/poor risk subjects.
All Intermediate/Poor Risk Treated Subjects	All intermediate/poor risk subjects who received any dose of study therapy. This population will be used for subset analyses of exposure and safety analyses on intermediate/poor risk subjects.
Pharmacokinetic Subjects	All subjects with available serum time-concentration data from randomized subjects dose with nivolumab and cabozantinib.
Immunogenicity Subjects	All subjects with available data from randomized subjects dose with nivolumab and cabozantinib.
PD-L1 Treated Subjects	All subjects with a PD-L1 assessment at baseline who received any dose of study therapy.

7 STATISTICAL ANALYSES

7.1 General Methods

Unless otherwise noted, discrete variables will be tabulated by the frequency and proportion of subjects falling into each category, grouped by treatment. Percentages given in these tables will be rounded to the first decimal and, therefore, may not always sum to 100%. Percentages less than 0.1 will be indicated as '< 0.1'. Continuous variables will be summarized by treatment group using the mean, standard deviation, median, minimum, and maximum values.

Time-to-event variables (e.g. time-to resolution) will be analyzed using the Kaplan-Meier technique. When specified, the median will be reported along with 95% CI using Brookmeyer and Crowley method[5] (using log-log transformation for constructing the confidence intervals[6]).

Unless otherwise specified, the stratified log-rank test will be performed to test the comparison between time to event distributions. Unless otherwise specified, the stratified hazard ratio between 2 groups along with CI will be obtained by fitting a stratified Cox model with the group variable as a unique covariate.

Confidence intervals for binomial proportions will be derived using the Clopper-Pearson method.[7] The unweighted difference in ORRs between the two treatment arms and corresponding asymptotic 95% CI will be estimated using a Newcombe method.[8]

P-values from sensitivity analyses for efficacy endpoints are for descriptive purpose only and there will be no multiplicity adjustment for these analyses.

The conventions to be used for imputing missing and partial dates for analyses requiring dates are described in Section 8.

Note that, the outputs will present Arm A, Arm B, and Arm C in the following way:

- All analysis specified in Section 7.3 and 7.4 will present A, B, and C together.
- For purposes of efficacy analyses, all randomized subjects in Arm A and Arm C will be tabulated and presented in the same table. A few seperate efficacy summary tables will be generated for those subjects randomized to Arm B, which will be specified in the corresponding sections.
- For purposes of safety analyses, all treated subjects in Arm A and Arm C will be tabulated and presented in the same table. A few separate safety summary tables will be generated for those subjects treated in Arm B, which will be specified in the corresponding sections.
- Summaries of outcome research and biomarker data will be tabulated on the Arm A and C subjects only.

7.1.1 Adverse Events, Serious Adverse Events, Multiple events, Select Adverse Events, Other Events of Special Interest and Immune-Mediated Adverse Events

Drug-related AEs are those events with relationship to study drug "Related", as recorded on the CRF. If the relationship to study drug is missing, the AE will be considered as drug-related.

Serious adverse events consist of AEs deemed serious by the Investigator and flagged accordingly in the CRF and clinical database.

Adverse events leading to study drug discontinuation are AEs with action taken regarding study drug(s) = "Drug was discontinued".

Adverse events leading to dose delay are AEs with action taken regarding study drug(s) = "Drug was delayed".

Adverse event that led to dose delay of the oral drug (similarly defined as dose omission or dose interruption) will be coded with action "Drug was interrupted".

Adverse events leading to dose reduction are AEs with action taken regarding study drug(s) = "Dose was reduced".

Adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA), and the most recent version of the dictionary at the time of the database lock will be used. Adverse events results will be graded for severity using NCI Common Terminology Criteria for Adverse Events (CTCAE) and the most recent version of the criteria at the time of the database lock will be used.

In the AE summary tables, unless otherwise specified, subjects will be counted only once at the Preferred Term (PT), only once at the System Organ Class (SOC), and only once at subject level for the counting of total number of subjects with an AE. The AE tables will be sorted by the SOCs and then PTs. SOC will be ordered by descending frequency overall and then alphabetically. PTs will be ordered within SOC by descending frequency overall and then alphabetically. The sorting will be done based on the 'Any Grade' column of the experimental arm when arms are presented side-by-side.

Unless otherwise specified, the AE summary tables will be restricted to on-treatment events regardless of the causality.

Analyses that take into account the multiple occurrences of a given adverse event will be conducted (see Section 7.6.9). To prepare these analyses, the CRF data will be processed according to standard BMS algorithms[9] in order to collapse adverse event records into unique records based on the preferred term. These data will be presented as the rate per 100 person-years of exposure. These analyses will take into account all on-treatment events (allowing more than 1 event per subject) and the total exposure time. The person-year exposure will be computed as the sum over the subjects' exposure expressed in years where the exposure time is defined as

- (Date of last dose of study treatment - date of first dose of study treatment + 31 days (or 101 days, depending on the analysis))/365.25, for subject who are off study treatment and were

followed for at least 30 days (or 100 days, depending on the analysis) after last dose of study treatment.

- (Last known alive date - date of first dose of study treatment +1)/365.25, for subjects who are still on-treatment or who are off study treatment and were followed less than 30 days (or 100 days depending on the analysis) after last dose of study treatment.

7.1.1.1 Select Adverse Events (EU/ROW Submissions)

The select Adverse Events (select AEs) consist of a list of preferred terms grouped by specific category (e.g. pulmonary events, gastrointestinal events categories, etc.). AEs that may differ from or be more severe than AEs caused by non-immunotherapies and AEs whose early recognition and management may mitigate severe toxicity are included as select AEs. Categories of select AEs may include subcategories (e.g. adrenal disorders, diabetes, pituitary disorders, and thyroid disorders are subcategories of the endocrine event category).

The list of MedDRA preferred terms used to identify select adverse events is revisited quarterly and updated accordingly. The preferred terms used for the selection at the time of the database lock will be provided by categories/subcategories.

In addition to the frequency and worst severity of select AEs, time-to onset, time-to resolution, and time-to resolution where immune modulating medication was initiated will be analyzed for each specific category/subcategory of drug-related select AEs when applicable.

Further details on the definitions time-to onset and time-to resolution are described in APPENDIX 1.

7.1.1.2 Other Events of Special Interest

Other events of special interest (OEOSI) consist of a list of preferred terms grouped by specific category (e.g. Myositis Event, Myocarditis Event, Demyelination Event, Guillain-Barre Syndrome, Pancreatitis Event, Uveitis Event, Encephalitis Event, Myasthenic Syndrome, Rhabdomyolysis Event, Graft Versus Host Disease). The list of MedDRA preferred terms used to identify OEOSI is revisited quarterly and updated accordingly. The preferred terms used for the selection at the time of the database lock by categories will be provided.

7.1.1.3 Immune-Mediated Adverse Events (US Submission)

In order to further characterize AEs of special clinical interest, analysis of immune-mediated AEs (IMAE) will be conducted. IMAEs are specific events (or groups of PTs describing specific events) that include pneumonitis, diarrhea/colitis, hepatitis, nephritis/renal dysfunction, rash, endocrine (adrenal insufficiency, hypothyroidism/thyroiditis, hypothyroidism, thyroiditis, hyperthyroidism, diabetes mellitus, and hypophysitis), and other specific events, considered as potential immune-mediated events by investigator that meet the definition summarized below:

- those occurring within 100 days of the last dose,
- regardless of causality,
- treated with immune-modulating medication (of note, endocrine AEs such as adrenal insufficiency, hypothyroidism/thyroiditis, hypothyroidism, thyroiditis, hyperthyroidism,

diabetes mellitus, and hypophysitis are considered IMAEs regardless of immune-modulating medication use, since endocrine drug reactions are often managed without immune-modulating medication).

- with no clear alternate etiology based on investigator assessment, or with an immune-mediated component

The list of MedDRA preferred terms used to identify IMAEs is revisited quarterly and updated accordingly. The preferred terms used for the selection at the time of the database lock by categories will be provided.

7.1.2 Laboratory Tests

Clinical laboratory parameters (hematology, serum chemistry and electrolytes) will be evaluated.

Laboratory tests will be graded using the NCI Common Terminology Criteria, and the most recent version of the criteria at the time of the database lock will be used.

Clinical laboratory data will be first analyzed using International System of Units (SI).

Analyses will be repeated using US conventional units.

In the laboratory summary tables, unless otherwise specified, subjects will be counted only once for each lab parameter according to their worst on treatment CTC grade (worst being the highest CTC grade). The laboratory tables and listings will be sorted by laboratory category, laboratory subcategory and laboratory test code sequence number.

7.1.3 Immunogenicity Data

Blood samples for immunogenicity analysis will be collected from subjects assigned to the experimental treatment group(s) according to the protocol schedule. Samples will be evaluated for development of Anti-Drug Antibody (ADA) by a validated electrochemiluminescent (ECL) immunoassay.

7.2 Study Conduct

The following programmable deviations will be considered as relevant protocol deviations and summarized by treatment group and overall in all randomized subjects. Non-programmable relevant eligibility and on-treatment protocol deviations, as well as significant (both programmable and non-programmable) eligibility and on-treatment protocol deviations will be reported through ClinSIGHT listings.

Eligibility:

- Subjects with baseline KPS < 70%
- Subjects who received prior systemic anti-cancer treatment in the metastatic setting
- Subjects without histologically confirmed RCC with a clear-cell component, documented advanced or metastatic RCC

On-study:

- Subjects receiving anti-cancer therapy (chemotherapy, hormonal therapy, immunotherapy, standard or investigational agents for treatment of cancer) while on study therapy
- Subjects treated differently than as randomized (subjects who received the wrong treatment, excluding the never treated)

Enrollment by country and site, and enrollment by month will be summarized and listed for all enrolled subjects.

A by-subject listing of batch numbers for all treated subjects will be provided.

7.3 Study Population

Analyses in this section will be tabulated for all randomized subjects by treatment group as randomized, unless otherwise specified.

7.3.1 Subject Disposition

The total number of subjects enrolled (randomized or not randomized) will be presented along with the reason for not being randomized. This analysis will be performed on the all enrolled subjects population only.

Number of subjects randomized but not treated along with the reason for not being treated will be tabulated by treatment group as randomized.

Number of subjects who discontinued study treatment along with corresponding reason will be tabulated by treatment group as treated. Reason for discontinuation will be derived from subject status CRF page. This analysis will be performed only on the all treated subjects population.

A by-subject listing for all treated subjects will be provided showing the subject's off treatment date and whether the subject continue in the study along with the reason for going off study. A by-subject listing for all enrolled subjects will also be provided, showing whether the subject was randomized along with the reason for not being randomized.

Note that for all intermediate/poor risk treated subjects in Arm A and Arm C, number of subjects who discontinued study treatment along with corresponding reason will be tabulated by treatment group as treated.

7.3.2 Demographics and Other Baseline Disease Characteristics

The following demographic and baseline disease characteristics will be summarized and listed by treatment group as randomized:

- Age
- Age categorization (< 65, ≥ 65 and < 75, ≥ 75 and < 85, ≥ 85, ≥ 75, ≥ 65)
- Sex (Male, Female)
- Race (White, Black or African American, Asian, Other)
- Region (Region (US/Canada/W.Europe/N.Europe vs. ROW) (source: IRT)

- Ethnicity (Hispanic/Latino and Not Hispanic/Latino)
- Karnofsky performance status (70, 80, 90, 100)
- Baseline IMDC prognostic score (0, 1-2, ≥ 3) (source: IRT)
- Baseline IMDC prognostic score (0, 1-2, ≥ 3) (source: CRF)
- Time from initial disease diagnosis to randomization (<1 year, ≥1 year)
- Baseline LDH level (≤ 1.5 x ULN, >1.5 x ULN)
- Hemoglobin (<LLN, ≥ LLN)
- Corrected Calcium (≤ 10 mg/dl, >10mg/dl)
- Absolute Neutrophil Count (≤ ULN, > ULN)
- Platelet Count (≤ ULN, > ULN)
- Baseline Alkaline phosphatase (< ULN, ≥ ULN)
- Prior nephrectomy (Yes, No)
- Prior radiotherapy (Yes, No)
- Baseline PD-L1+ status based on a 1% cut off (≥ 1% vs. < 1% or indeterminate)
- Baseline PD-L1+ status based on a 5% cut off (≥ 5% vs. < 5% or indeterminate)
- Baseline PD-L1+ status based on a 10% cut off (≥ 10% vs. < 10% or indeterminate)
- Sites of diseases (all lesions)
- Number of disease sites per subject (1, 2, 3, 4 , >4)
- Tumor burden: sum of the diameters of target lesions at baseline
- Most common sites of metastasis (lung, lymph node, bone, liver, adrenal gland)
- Sarcomatoid features (Yes, No)
- Stage at the initial diagnosis (Stage IV, non-Stage IV)
- Pre-treatment events tumor assessment (per Investigator)

Summary table (cross-tabulation) by treatment group for stratification factor (except for region) will be provided to show any discrepancies between what was reported through IRT vs. other data sources at baseline. This summary will be performed based on all randomized subjects.

- IMDC Prognostic Score: 0 versus 1-2 versus 3-6
- Region: US/Canada/W Europe/N Europe versus ROW
- PD-L1 tumor expression: ≥ 1% versus < 1% or indeterminate

A listing of randomization scheme presenting randomized treatment group and as treated treatment group will be provided for all randomized subjects.

Note that for all intermediate/poor risk subjects in Arm A and Arm C, demographic and baseline disease characteristics will be summarized and listed by treatment group as randomized.

7.3.3 Medical History

A by-subject listing of general medical history for all randomized subjects will be provided.

7.3.4 Prior Therapy Agents

Prior adjuvant or neo-adjuvant therapy will be summarized by treatment group and overall.

Prior systemic cancer therapy will be summarized by treatment group and overall and listed by subject.

Prior radiotherapy and prior surgery related to cancer will be listed by subject.

7.3.5 Physical Examinations

Subjects with abnormal baseline physical examination will be listed by subject.

7.3.6 Baseline Physical Measurements

Baseline physical measurements will be listed by subject.

7.4 Extent of Exposure

Listings will include all available exposure data. Analyses will be performed by treatment group "as treated" in all treated subjects, unless otherwise specified.

7.4.1 Administration of Study Therapy

The following parameters will be summarized (descriptive statistics) by study therapy and treatment group:

- Number of doses received
- Cumulative dose
- Relative dose intensity (%) using the following categories: < 50%; 50 - < 70%; 70 - < 90%; 90 - < 110%; ≥ 110%
- Average daily dose

Duration of study therapy will be summarized (descriptive statistics) by treatment group.

A by-subject listing of dosing of study medication (record of study medication, infusion details, and dose changes) will be also provided.

Note that similar study therapy table will be summarized for all intermediate/poor risk treated subjects in Arm A and Arm C.

Table 7.4.1-1: Study Therapy Parameter Definitions for Arm A and C

	Nivolumab	Cabozantinib	Sunitinib
Dosing Schedule per Protocol	240 mg every 2 weeks	40 mg PO once daily	50 mg PO once daily for 4 weeks followed by 2 weeks off.
Dose	mg	mg	mg
Cumulative Dose	mg sum of the doses administered to a subject	mg sum of the doses administered to a subject	mg sum of the doses administered to a subject
Relative Dose Intensity (%)	[Cum dose (mg) /((Last dose date– First dose date + 14) × 240/14)] × 100	See below	See below
Duration of Treatment	*Last dose date - Start dose date +1*	*Last dose date - Start dose date + 1*	*Last dose date - Start dose date +15*

Additional Parameters -Cabozantinib treatment

Average daily dose (in mg/day) is defined as:

Sum of all Cabozantinib doses in mg actually received / duration of treatment in days.

Since Cabozantinib treatment consists of 40 mg PO daily dose, the planned dose intensity of Cabozantinib is 40 mg/day.

Relative dose intensity for Cabozantinib (%) is defined as: *(Average daily dose / 40) x 100.*

Additional Parameters -Sunitinib treatment

Average daily dose (in mg/day) is defined as:

Sum of all Sunitinib doses in mg actually received / duration of treatment in days

Since Sunitinib treatment consists of 50 mg PO daily dose for 4 weeks followed by 2 weeks of washout period, the planned dose intensity of Sunitinib is 33.33 mg/day (50 mg x 28 days / 42 days).

Relative dose intensity for Sunitinib (%) is defined as: *(Average daily dose / 33.33) x 100.*

Table 7.4.1-2: Study Therapy Parameter Definitions for Arm B (Cycle 1-4)

	Nivolumab	Ipilimumab	Cabozantinib
Dosing Schedule per Protocol	3 mg/kg every 3 weeks for 4 doses	1 mg/kg every 3 weeks for 4 doses	40 mg PO once daily

Table 7.4.1-2: Study Therapy Parameter Definitions for Arm B (Cycle 1-4)

	Nivolumab	Ipilimumab	Cabozantinib
Dose	Dose (mg/kg) is defined as Total Dose administered (mg)/Most recent weight (kg). Dose administered in mg at each dosing date and weight are collected on the CRF.	Dose (mg/kg) is defined as Total Dose administered (mg)/Most recent weight (kg). Dose administered in mg at each dosing date and weight are collected on the CRF.	mg
Cumulative Dose	Cum Dose (mg/kg) is the sum of the doses administered to a subject.	Cum Dose (mg/kg) is the sum of the doses administered to a subject.	mg sum of the doses administered to a subject
Cycle Duration$_{(i)}$ (wk)	(Dose date$_{(i+1)}$ - Dose date$_{(i)}$)/7	(Dose date$_{(i+1)}$ - Dose date$_{(i)}$)/7	N/A
Cycle Intensity$_{(i)}$ (mg/kg/wk)	Dose$_{(i)}$/Cycle Duration$_{(i)}$	Dose$_{(i)}$/Cycle Duration$_{(i)}$	N/A
Relative Cycle Intensity $_{(i)}$ (%)	(Cycle Intensity$_{(i)}$/intended dose per week)(i) * 100	(Cycle Intensity$_{(i)}$/intended dose per week)(i) * 100	N/A
Relative Dose Intensity (%)	Sum of all Relative Cycle Intensities divided by N	Sum of all Relative Cycle Intensities divided by N	(Average daily dose / 40) x 100
Duration of Treatment	*Last dose date - Start dose date +1*	*Last dose date - Start dose date +1*	*Last dose date - Start dose date +1*

Table 7.4.1-3: Study Therapy Parameter Definitions for Arm B (Cycle 5 Onward)

	Nivolumab	Cabozantinib
Dosing Schedule per Protocol	240 mg every 2 weeks	40 mg PO once daily
Dose	mg	mg
Cumulative Dose	mg sum of the doses administered to a subject	mg sum of the doses administered to a subject
Relative Dose Intensity (%)	[Cum dose (mg) /((Last dose date– First dose date + 14) × 240/14)] × 100	(Average daily dose / 40) x 100
Duration of Treatment	*Last dose date - Start dose date +1*	*Last dose date - Start dose date + 1*

7.4.2 Modifications of Study Therapy

7.4.2.1 Dose Delays

Each nivolumab infusion or sunitinib dose may be delayed. A dose will be considered as actually delayed if the delay is exceeding 3 days (ie greater than or equal to 4 days from scheduled dosing date) for nivolumab. Reason for dose delay will be retrieved from CRF dosing pages. It is worth noting that during the two week mandatory washout period for sunitinib, a daily dose of 0 mg will be entered in the CRF pages, with corresponding reason for dose modification recorded as "No Change". For cabozantinib, a daily dose of 0 mg entered in the CRF pages will be considered as delay. If cabozantinib is given every other day, then a daily dose of 0 mg will be entered in the CRF pages, with corresponding reason for dose modification recorded as "No Change".

The following parameters will be summarized by treatment group:

- Number of subjects with at least one dose delayed, the number of dose delays per subject, the reason for dose delay and the length of dose delay.

Note that similar dose delay summary table for Arm A and Arm C will be summarized for all intermediate/poor risk treated subjects.

For Arm B, both nivolumab and ipilimumab can be delayed at the same cycle. A dose will be considered as actually delayed if the delay is exceeding 3 days (ie greater than or equal to 4 days from scheduled dosing date) for nivolumab and ipimumab. Cabozantinib a daily dose of 0 mg entered in the CRF pages will be considered as delay. Similar table will be produced for Arm B subjects previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinical planned events, per protocol.

7.4.2.2 Infusion Interruptions and Rate Changes

Each nivolumab or ipilimumab infusion can be interrupted and/or the IV infusion rate can be reduced. This information will be retrieved from CRF dosing pages.

The following parameters will be summarized by treatment group:

- Number of subjects with at least one dose infusion interruption, the reason for interruption, and the number of infusion interruptions per subject.
- Number of subjects with at least one IV infusion rate reduction, the reason for reduction and the number of infusion with IV rate reduction per subject.

Note that similar summary table for Arm A and Arm C will be summarized for all intermediate/poor risk treated subjects.

7.4.2.3 Dose Escalations

Dose escalations are permitted for cabozantinib and sunitinib but not for nivolumab and ipilimumab.

7.4.2.4 Dose Reductions

Dose reductions are permitted for cabozantinib and sunitinib but not for nivolumab and ipilimumab.

Dose reduction for subjects treated with sunitinib is defined as at least one day with a non zero dose smaller than 50 mg and smaller than previous non zero dose with a CRF reason different from "Dosing Error" or "No Change".

Dose reduction for subjects treated with cabozantinib is defined as at least one day with a non zero dose smaller than 40 mg and smaller than previous non zero dose with a CRF reason "Dosing Error" or "No Change".

Note that similar dose redcution summary table for Arm A and Arm C will be summarized for all intermediate/poor risk treated subjects.

The following summaries will be presented for the cabozantinib component of study treatment:

i. For dose reductions due to AE

   Categorical summaries for:

   - Subjects with any dose reduction
   - Dose levels received by a subject
   - Lowest non-zero dose level received
   - Last non-zero dose level received
   - Last dose level received (including dose holds)

   Descriptive statistics for:

   - Duration of treatment in months for each dose level (40 mg, 20 mg, 0 mg)
   - Time to second dose level reduction (first receipt of 20mg) (days)

ii. Summaries for dose holds due to AE (those with 0 mg due to AE):
   - Descriptive statistics for number of dose holds due to an AE
   - Descriptive statistics for duration of dose holds per dose hold and per subject due to an AE, calculated as (stop date of hold – start date of hold + 1)
   - Categorical summary for subjects with duration of holds due to an AE that can be classified as any number of days, ≥ 7 days, ≥ 14 days, ≥ 21 days, and >42 days
   - Descriptive statistics for time to first dose hold, time to first dose hold that ≥ 7 days, ≥ 14 days, ≥ 21 days, and >42 days. The time to dose hold is calculated as (start date of the hold – first dose date + 1)

- Descriptive statistics for time to second dose hold, time to second dose hold that was $\geq$ 7 days, $\geq$ 14 days, $\geq$ 21 days, and >42 days.

iii. Summaries for dose modifications (defined as a reduction or hold) due to AE:
- Frequency counts and percentages for subjects with any dose modifications
- Descriptive statistics for number of dose modifications (0-3)
- Descriptive statistics for time to the first dose modification
- Descriptive statistics for time to the second dose modification

7.4.2.5 Dose Omissions

Dose omissions are not permitted.

7.4.3 Concomitant Medications

Concomitant medications, defined as medications other than study medications which are taken at any time on-treatment (i.e. on or after the first day of study therapy and within 100 days following the last dose of study therapy), will be coded using the UMC WHO Drug Global Dictionary.

The following summary table will be provided:

- Concomitant medications (subjects with any concomitant medication, subjects by medication class and generic term)

A by-subject listing will accompany the table.

7.4.3.1 Immune Modulating Medication

Immune modulating concomitant medications are medications entered on an immune modulating medication form or available from the most current pre-defined list of immune modulating medications. The list of anatomic class, therapeutic class and generic name used for the selection at the time of the database lock will be provided.

The percentage of subjects who received immune modulating concomitant medication for

- management of adverse event
- premedication
- other use
- any use
- management of drug-related select adverse event (any grade, grade 3-5) by select AE category/ subcategory (EU/ROW Submissions)
- management of IMAEs (any grade, grade 3-5) by IMAE category (US Submission)

will be reported separately for each treatment group (percentages of treated subjects by medication class and generic term).

For each category/subcategory of drug-related select AEs (any grade, grade 3-5) and IMAEs (any grade, grade 3-5), the following will be reported for each treatment group:

- The total immune modulating medication treatment duration (excluding overlaps), duration of high dose of corticosteroid, initial dose of corticosteroid, and tapering duration (summary statistics)

Duration represents the total duration the subject received the concomitant medication of interest. If the subject took the medication periodically, then DURATION in the summation of all use. Initial dose represents the dose of the concomitant medication of interest received at the start of the event. In the case multiple medications started on the same date, the highest equivalent dose is chosen and converted to mg/kg by dividing by the subject's recent weight.

These analyses, except the ones related to IMAEs will be conducted using the 30-day safety window. The analyses related to IMAEs will be conducted using the 100-day safety window.

7.4.3.2 Subsequent Cancer Therapy

Number and percentage of subjects receiving subsequent cancer therapies will be summarized for all randomized subjects. Categories include:

- Subsequent systemic therapy
- Subsequent surgery for treatment of tumors
- Subsequent radiotherapy for treatment of tumors

A by-subject listing of subsequent cancer therapy will also be produced for all randomized subjects.

Note that similar tables will be summarized for all intermediate/poor risk subjects in Arm A and Arm C.

7.5 Efficacy

Analyses in this section will be tabulated for all randomized subjects in Arm A and Arm C, unless otherwise specified. A few seperate efficacy summary tables will be generated for those subjects who randomized to Arm B prior to Revised Protocol 01, which is specified in the relevant subsections below.

Principal analyses of progression free survival (PFS) and objective response rate (ORR) will be based on the Blinded Independent Central Review (BICR) evaluation, unless noted otherwise.

Analyses in this section will be tabulated for all randomized subjects by treatment group as randomized, unless otherwise specified.

Unless stated otherwise, whenever a stratified analysis is specified, the following stratifications factors (recorded at randomization as per IRT) will be used:

- IMDC Prognostic Score: 0 versus 1-2 versus 3-6
- Region: US/Canada/W Europe/N Europe versus ROW
- PD-L1 tumor expression: $\geq$ 1% versus < 1% or indeterminate

The key secondary objective Overall Survival among all randomized subjects will be tested after conducting the primary objective analyses of PFS on all randomized subjects. For assessing this secondary objective of this study, a hierarchical testing procedure[10] will be used so that the overall experiment-wise Type I error rate is two-sided 0.05.

Confidence intervals (CI) for primary and secondary endpoint analyses included in hierarchy will be based on nominal significance level adjusted for primary endpoints and interim analyses to preserve overall type one error rate.

Alpha (α) for the CI will be the same as nominal significance level for hypothesis testing. CIs for other endpoints will be at the two-sided 95% level. All p-values reported will be two-sided. P-values will be rounded to the fourth decimal place. Point estimates and confidence bounds for efficacy variables will be rounded to the second decimal place.

A by-subject listing of efficacy results will be presented including treatment group, treatment duration, BICR progression date, overall survival, death date, etc.

7.5.1 Analysis of Progression-Free Survival

The primary objective of the study is to compare the PFS per BICR of Arm A to Arm C in randomized subjects with previously untreated (first line) advanced or metastatic RCC. All the analyses outlined in this section are specified for the all randomized subjects population in Arm A and Arm C, unless otherwise specified.

PFS per BICR will be compared between the treatment groups via stratified log-rank test among all randomized subjects at a two-sided $\alpha = 0.05$ level. The stratification factors will be IMDC prognostic risk score (0 vs 1-2 vs 3-6), region (US/Canada/W Europe/N Europe vs ROW)' and PD-L1 status ($\geq$ 1% vs < 1% or indeterminate).

The primary definition of PFS adjusting for subsequent anticancer therapy will be used in this analysis. The two-sided log-rank p-value will be reported.

The estimate of the PFS hazard ratio between treatment groups will be calculated using a stratified Cox proportional hazards model, with treatment as the sole covariate. Ties will be handled using the exact method. A two-sided 95% CI for the hazard ratio will also be presented.

The PFS function for each treatment group will be estimated using the KM product limit method and will be displayed graphically. A two-sided 95% CI for median PFS in each treatment group will be computed via the log-log transformation method. PFS rates at fixed time points (e.g. 6 months, depending on the minimum follow-up) will be presented along with their associated 95% CIs. These estimates will be derived from the Kaplan Meier estimate and corresponding CIs will be derived based on Greenwood[11] formula for variance derivation and on log-log transformation applied on the survivor function[12].

Analyses of PFS will also be conducted based on the secondary definition of PFS. These analyses will be the same as those specified above.

The source of PFS event (progression or death) will be summarized by treatment group. The status of subjects who are censored (as per primary definition of PFS) in the PFS KM analysis will be tabulated for each treatment group including the following categories:

- On-study (on-treatment, in follow-up)
- Off-study (lost to follow-up, withdraw consent, never treated)
- No baseline tumor assessment
- No on-study tumor assessment and no death
- Received subsequent anticancer therapy

A by-subject listing will be presented including treatment group, PFS duration under the primary definition, PFS duration on the ITT definition, whether the subject was censored under the primary definition, and if censored, the reason, and whether the subject was censored under the ITT definition, and if censored, the reason.

A by-subject listing of lesion evaluations per BICR will be presented.

Note that similar tables (KM plot, median PFS with its 95% CI, PFS rates, source of PFS event, and status of subjects who are censored) along with by-subject listing will be summarized for all intermediate/poor risk subjects in Arm A and Arm C.

For those subjects who randomized to Arm B prior to Revised Protocol 01, the PFS function will be estimated using the KM product limit method and reported seperately for Arm B. A two-sided 95% CI for median PFS will be computed via the log-log transformation method. A by-subject listing will be presented including treatment group, PFS duration under the primary definition, PFS duration on the ITT definition, whether the subject was censored under the primary definition, and if censored, the reason, and whether the subject was censored under the ITT definition, and if censored, the reason.

7.5.2 Supportive Analyses of Progression-Free Survival

The following sensitivity analyses will be conducted using both the primary and the secondary definition of PFS in all randomized subjects:

1. Delayed effect of immunotherapy interventions may cause a late separation in the PFS KM curves and non-proportional hazards. PFS (as determined by BICR) will be compared between treatment groups via two-sided 0.05 stratified weighted log-rank test among subjects. The primary definition of PFS will be used in this analysis. The two-sided stratified weighted log-rank p-value will be reported using G (rho = 0, gamma = 1) weights, in the terminology of Fleming and Harrington[13].

   The Fleming Harrington test can be unstable, so it is possible, though uncommon, that thep-value for this trial will not be estimable.

   The estimate of the PFS hazard ratio in the period before and following 6 months will be calculated using a stratified time-dependent Cox model with effects for treatment and period-by-treatment interaction. In this model, period is a binary variable indicating pre-

vs. post-6 months. Ties will be handled using the exact method. A two-sided 95% CI for the hazard ratio will also be presented.

2. A multivariate Cox regression model will be used in order to estimate the treatment effect after adjustment for possible imbalances in known or potential prognostic factors. The factors used in the randomization, which, by definition, will be balanced across treatment groups, will still be included in the model as stratification factors. However, all additional factors will be incorporated as covariates. The additional factors, which are all measured at baseline, will include:

    a. Age categorization (< 65 vs. ≥ 65 - < 75 vs. ≥ 75)
    b. Gender (Male vs. Female)
    c. Race (White, Black or African American, Asian, Other)
    d. Region (US/Canada/W.Europe/N.Europe vs. ROW)
    e. IMDC score (0 vs 1-2 vs 3-6)
    f. Karnofsky performance status (70, 80, 90, 100)
    g. Prior Nephrectomy (Yes, No)
    h. LDH level (≤ 1.5 x ULN , > 1.5 x ULN)
    i. Baseline PD-L1+ status based on a 1% cut off
    j. Number of disease site (1 vs. 2 vs. >2)

    The level of the covariate normally associated with the worst prognosis will be coded as the reference level. The hazard ratio associated with treatment and with each of the baseline covariates will be presented along with associated 95% CIs.
3. PFS using stratification factors as obtained from the baseline CRF pages (instead of IRT). The hazard ratio associated with treatment will be presented along with the associated two-sided 95% CIs. This analysis will be performed only if at least one stratification variable/factor at randomization (as per IRT) and baseline are not concordant for at least 10% of the randomized subjects.
4. PFS using the investigator's assessment. The hazard ratio associated with treatment and median PFS will be presented along with the associated two-sided 95% CIs.

    A cross tabulation of PFS assessment by BICR versus PFS assessment by investigator will be presented, by treatment group. Concordance Rate of event will be computed as the frequency with which investigator and BICR agree on classification of a subject as event versus censored as a proportion of the total number of randomized subjects assessed by both the investigator and BICR.

    A by-subject listing of PFS assessment per BICR and investigator will be presented.
5. PFS using an un-stratified log rank test. The hazard ratio associated with treatment will be presented along with the associated two-sided 95% CIs.

6. PFS using an un-stratified Cox proportional hazards model, adjusted, using as covariates only the stratification factors used in randomization. The hazard ratio associated with treatment will be presented along with the associated two-sided 95% CIs.
7. PFS for subjects with no relevant protocol deviations. This analysis will be conducted only if there are more than 10% subjects with relevant protocol deviations. The hazard ratio associated with treatment will be presented along with the associated two-sided 95% CIs.
8. The method of Gail and Simon[14] will be used to test for a qualitative interaction between treatment and strata. This test will be conducted at $\alpha = 0.10$ level. The p-value reported from this specific analysis is for descriptive purposes alone.
9. To examine the assumption of proportional hazards in the Cox regression model, in addition to treatment, a time-dependent variable defined by treatment by time interaction will be added into the model. A two-sided Wald Chi-square p-value of less than 0.1 may indicate a potential nonconstant treatment effect. In that case, additional exploratory analyses may be performed.

7.5.3 Subset Analyses of Progression-Free Survival

The influence of baseline and demographic characteristics on the treatment effect among all randomized subjects will be explored via exploratory subset analyses. The median PFS based on KM product-limit method along with two-sided 95% CIs will be produced for the following subgroups:

- Age
- Age categorization (< 65, ≥ 65 and < 75, ≥ 75 and < 85, ≥ 85, ≥ 75, ≥ 65)
- Sex (Male, Female)
- Race (White, Black or African American, Asian, Other)
- Region (Region (US/Canada/W.Europe/N.Europe vs. ROW) (source: IRT)
- Ethnicity (Hispanic/Latino and Not Hispanic/Latino)
- Karnofsky performance status (100-90, 80-70)
- Baseline IMDC prognostic score (0, 1-2, ≥ 3) (source: IRT)
- Baseline IMDC prognostic score (0, 1-2, ≥ 3) (source: CRF)
- Time from initial disease diagnosis to randomization (<1 year, ≥1 year)
- Baseline LDH level (≤ 1.5 x ULN, >1.5 x ULN)
- Hemoglobin (<LLN, ≥ LLN)
- Corrected Calcium (≤ 10 mg/dl, >10mg/dl)
- Absolute Neutrophil Count (≤ ULN, > ULN)
- Platelet Count (≤ ULN, > ULN)
- Baseline Alkaline phosphatase (< ULN, ≥ ULN)
- Prior nephrectomy (Yes, No)
- Prior radiotherapy (Yes, No)

- Bone metastasis (Yes, No)
- Sarcomatoid features (Yes, No)
- Stage at the initial diagnosis (Stage IV, non-Stage IV)
- Baseline PD-L1+ status based on a 1% cut off (≥ 1% vs. < 1% or indeterminate)
- Baseline PD-L1+ status based on a 5% cut off (≥ 5% vs. < 5% or indeterminate)
- Baseline PD-L1+ status based on a 10% cut off (≥ 10% vs. < 10% or indeterminate)

A forest plot of the unstratified PFS hazard ratios (along with the 95% CIs) will be produced for each level of the subgroups listed above. The analysis comparing treatment (i.e., Hazard Ratio) will be conducted if the number of subjects in the subgroup category is more than 10.

7.5.4 Analysis of Overall Survival

One of the secondary objectives of the study is to compare the overall survival of Arm A to Arm C in all randomized subjects. All the analyses outlined in this section are specified for the all randomized subjects population in Arm A and Arm C, unless otherwise specified.

If the formal analysis of PFS among all randomized subjects is statistically significant, the formal interim analysis of OS among all randomized subjects will be tested, as per hierarchical testing procedure.

Overall survival will be compared between the treatment groups at the interim and final analyses, using stratified log-rank test. The stratification factors will be those used in the analysis of PFS. An O'Brien and Fleming α-spending function will be employed to determine the nominal significance levels for the interim and final analyses. The stratified hazard ratio between the treatment groups will be presented along with 100*(1- α)% CI (adjusted for interim). In addition, two-sided p-value will also be reported for the analysis of OS.

OS will be estimated using the KM techniques. A two-sided 95% CI for median OS in each treatment group will be computed via the log-log transformation method. OS rates at fixed time points (e.g. 6 months, depending on the minimum follow-up) will be presented along with their associated 95% CIs. These estimates will be derived from the Kaplan Meier estimate and corresponding CIs will be derived based on Greenwood formula for variance derivation and on log-log transformation applied on the survivor function.

The status of subjects who are censored in the OS KM analysis will be tabulated for each treatment group using the following categories:

- On-study (on-treatment, in follow-up)
- Off-study (lost to follow-up, withdraw consent, never treated)

A by-subject listing will be presented including treatment group, first and last dose date, whether the subject died, and if censored, the reason, event/censored date and OS duration.

Note that similar tables (KM plot, median OS with its 95% CI, OS rates, and status of subjects who are censored) along with by-subject listing will be summarized for all intermediate/poor risk subjects in Arm A and Arm C.

The analysis performed for PFS (detailed in section 7.5.1) for those subjects who randomized to Arm B prior to Revised Protocol 01 will be repeated for OS.

7.5.5 Subset Analysis of Overall Survival

The influence of baseline and demographic characteristics on the treatment effect among all randomized subjects will be explored via exploratory subset analyses. The median OS based on KM product-limit method along with two-sided 95% CIs will be produced for the same subgroups as used for PFS (see Section 7.5.3).

A forest plot of the unstratified OS hazard ratios (along with the 95% CIs) will be produced for each level of the subgroups listed above.

An analysis will be conducted if the number of subjects in the subgroup category is more than 10.

7.5.6 Current Status of PFS and OS Follow-up

The extent of follow-up for survival, defined as the time between randomization date and last known alive date (for subjects who are alive) or death date (for subjects who died), will be summarized descriptively (median, min, max, etc.) in months for all randomized subjects.

The currentness of follow-up for survival, defined as the time between last OS contact (i.e., last known alive date or death date) and cutoff date (defined by last subject last visit date), will be summarized in months for all randomized subjects. Subjects who died and subjects with last known alive date on or after data cut-off date will have zero value for currentness of follow-up.

Minimum follow-up of OS for all randomized subjects, defined as the time from cutoff date to last subject's randomization date, will be summarized in months.

Time from last evaluable tumor assessment to cutoff date in months will be summarized by treatment group and overall for all randomized subjects. Subjects who have a PFS event will be considered as current for this analysis. The secondary definition of PFS will be used for this summary.

In addition, time to treatment discontinuation will be summarized and presented by treatment group using a Kaplan-Meier curve whereby the last dose date will be the event date for those subjects who are off study therapy. Median duration of study therapy and associated 95% CI will be provided. Subjects who are still on study therapy will be censored on their last dose date.

A by-subject listing will also be produced to accompany the subject time from last evaluable tumor assessment.

7.5.7 Interim Analysis of Overall Survival

An independent statistician external to BMS will perform the analysis. In addition to the formal planned interim analyses for OS, the Data Monitoring Committee (DMC) will have access to

periodic un-blinded interim reports of efficacy and safety to allow a risk/benefit assessment. Details are included in the DMC charter.

Two interim analyses of OS are planned for this study. The first interim analysis of OS is planned at the time of final PFS analysis and expected after observing 165 deaths (approximately 65% of the targeted OS events) have been observed among all randomized subjects in Arm A and Arm C based on above accrual rate and the exponential distribution in each arm. These formal comparisons of OS will allow for early stopping for superiority, and the boundaries for declaring superiority will be derived based on the actual number of deaths using Lan-DeMets spending function with O'Brien and Fleming type of boundary in EAST version 6. If the first interim analysis is performed exactly at 165 deaths, the boundary in terms of statistical significance for declaring superiority would be 0.011 (HR=0.673 with 16 months improvement in median OS for the Arm A versus Arm C comparison (33 versus 49 months)). The second interim analysis of OS is expected after observing 211 deaths (approximately 83% of the targeted OS events) have been observed among all randomized subjects based on above accrual rate and the exponential distribution in each arm. The boundary for declaring superiority in terms of statistical significance for the second interim analysis after 211 events would be 0.025 (HR=0.734 with 12 months improvement in median OS for the doublet versus sunitinib comparison (33 versus 45 months). The boundary for declaring superiority in terms of statistical significance for the final analysis after 254 events would be 0.041 (HR=0.774 with 9.6 months improvement in median OS for the Arm A versus Arm C comparison (33 versus 42.6 months).

Note that if the analysis of PFS final analysis and first interim analysis OS is trigger with 8-months minimum follow-up on all randomized subjects, minimum 283 PFS events and 149 OS events, then the details of the first interim analysis of OS will be as follows:

- If the first interim analysis is performed exactly at 149 deaths, the boundary in terms of statistical significance for declaring superiority would be 0.007 (HR=0.643 with 18.3 months improvement in median OS for the Arm A versus Arm C comparison (33 versus 51.3 months)).

The DMC will review the safety and efficacy data from the informal interim analyses, BMS will remain blinded to these interim results and and will determine if the study should continue with or without changes or if accrual should be stopped. Subject enrollment will continue while waiting for the DMC's decisions.

The chair of the DMC and the sponsor can call an unscheduled review of the safety data.

If the formal analysis of PFS among all randomized subjects is statistically significant, the formal interim analysis of OS among all randomized subjects will be tested, as per hierarchical testing procedure.

At the time of the formal interim analysis for superiority of OS, the DMC may recommend continuing or stopping the trial. If the trial continues beyond the formal interim analysis, BMS will remain blinded to these interim results and the nominal critical point for the final OS analysis will be determined using the recalculated information fraction at the time of the interim analysis, as

described above. The final OS hazard ratio and corresponding confidence interval will be reported whereby the confidence interval will be adjusted accordingly (i.e. using the recalculated nominal α level at the final analysis).

If the trial is stopped for superiority of OS at the interim, the p-value from the interim stratified log-rank test will be considered the final primary analysis result.

7.5.8 Analysis of Objective Response Rate

One of the secondary objectives of the study is to evaluate the objective response rate in all randomized subjects in Arm A and Arm C. All the analyses outlined in this section are specified for the all randomized subjects population in Arm A and Arm C, unless otherwise specified.

The number and percentage of subjects in each category of BOR per BICR (complete response [CR], partial response [PR], stable disease [SD], progressive disease [PD], or unable to determine [UTD]) will be presented, by treatment group. Estimates of response rate, along with its exact two-sided 95% CI by Clopper and Pearson[15] will be presented, by treatment group.

Similar analyses will be repeated based on the investigator's assessment of ORR. A cross tabulation of BICR best response versus the investigator best response will be presented, by treatment group and by response categories. Concordance Rate of Responders will be computed as the frequency with which investigator and BICR agree on classification of a subject as responder vs. non responder/UTD as a proportion of the total number of randomized subjects assessed by both the investigator and BICR.

The following subject-level graphics will also be provided:

- For the responders only, time courses of the following events of interest will be graphically displayed: tumor response, progression, last dose received, and death.
- For response evaluable subjects (randomized subjects with baseline and at least one on-study tumor assessment),
  - A bar plot showing the best % reduction from baseline in sum of diameter of target lesions based on BICR assessment for each subject will be produced (excluding assessments after PD and assessments after start of subsequent anti-cancer therapy).
  - A plot of individual time course of tumor burden change per BICR assessment will be produced.

A by-subject listing of best overall response will be presented including treatment group, best overall response per BICR and dates of CR/PR/progression.

A by-subject listing of per time point tumor response per BICR will be presented.

Note that similar tables along with by-subject listing will be summarized for all intermediate/poor risk subjects in Arm A and Arm C.

For those subjects who randomized to Arm B prior to Revised Protocol 01, estimates of response rate, along with its exact two-sided 95% CI by Clopper-Pearson method, will be computed per BICR and investigator. DOR and TTR will also be evaluated per BICR.

7.5.9 Subset Analyses of Objective Response

The influence of baseline and demographic characteristics on the treatment effect will be explored via exploratory subset analysis. The subsets wil be same as those analyzed for PFS and will be reported based on the BICR assessment of ORR.

A forest plot of treatment effect on ORR per BICR in the above subgroups will be produced. The un-weighted differences in ORR between the two treatment groups and corresponding 95% two-sided CI using the method of Newcombe will be provided if the number of subjects in the subgroup category is more than 10.

7.5.10 Time to Tumor Response and Duration of Response

The analyses specified in this section will be conducted for all treatment arms. Duration of response (DOR) and time to response (TTR) will also be evaluated for subjects who achieved confirmed PR or CR. The DOR for each treatment group will be estimated using the Kaplan-Meier (KM) product limit method and will be displayed graphically. A table will be produced presenting number of events, number of subjects involved, medians, and 95% CIs for the medians. Median values of DOR, along with two-sided 95% CI in each treatment group will be computed based on a log-log transformation method.

The status of subjects who are censored in the DOR KM analysis will be tabulated for each treatment group including the following categories:

- Ongoing follow-up (current [last scan within adequate window vs cutoff date], not current)
- Off-study (lost to follow-up, withdraw consent, never treated)
- Received subsequent anticancer therapy.

TTR, which does not involve censoring, will be summarized by treatment group in all responders using descriptive statistics.

Cumulative Response Rates will be tabulated for Week 8, Month 4, 6, 8, and 12, and overall response rate will be provided.

A by-subject listing will be presented including treatment group, best response, time to response, duration of response, whether the subject was censored for duration of response, and, if so, the reason.

7.5.11 PFS2

One of the exploratory objectives of the study is to compare PFS2 between treatment groups in all randomized subjects.

PFS2 will be analyzed similarly to PFS:

- Median values based on KM method, along with two-sided 95% CI using Brookmeyer and Crowley method will be calculated. The estimate of standard error will be calculated using the Greenwood formula;
- PFS2 will be graphically displayed along with the median and 95% CI.

A by-subject listing of PFS and PFS2 will be provided.

7.6 Safety

Analyses in this section will be tabulated for all treated subjects by treatment group as treated, unless otherwise specified.

Analyses in this section will be tabulated for all treated subjects in Arm A and Arm C, unless otherwise specified. Limited selection of the summary tables will be generated for those subjects who randomized to Arm B prior to Revised Protocol 01, which is specified in the relevant subsections below.

7.6.1 Deaths

Deaths will be summarized by treatment group:

- All deaths, reasons for death.
- Deaths within 30 days of last dose received, reasons for death.
- Deaths within 100 days of last dose received, reasons for death.

A by-subject listing of deaths will be provided for the all enrolled subjects population.

Note that for all intermediate/poor risk treated subjects in Arm A and Arm C, deaths will be summarized by treatment group.

Similar tables will be presented for those subjects treated in Arm B.

7.6.2 Serious Adverse Events

Serious adverse events will be summarized by treatment group:

- Overall summary of SAEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.
- Overall summary of drug-related SAEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.

All analyses will be conducted using the 30-day safety window.
A by-subject SAE listing will be provided for the “enrolled subjects” population.

Note that for all intermediate/poor risk treated subjects in Arm A and Arm C, serious adverse events will be summarized by treatment group.

Similar tables will be presented for those subjects treated in Arm B.

7.6.3 Adverse Events Leading to Discontinuation of Study Therapy

AEs leading to discontinuation will be summarized by treatment group:

- Overall summary of AEs leading to discontinuation by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.
- Overall summary of drug-related AEs leading to discontinuation by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.

The analyses will be conducted using the 30-day safety window.

A by-subject AEs leading to discontinuation listing will be provided.

Note that for all intermediate/poor risk treated subjects in Arm A and Arm C, AEs leading to discontinuation will be summarized by treatment group.

Similar table will be presented for those subjects treated in Arm B.

Note that

- for Arm A, AEs leading to discontinuation from cabozantinib only, nivolumab only, and both cabozantinib and nivolumab.

- for Arm B, AEs leading to discontinuation from cabozantinib only, nivolumab and ipilimumab only, and from cabozantinib, nivolumab, and ipilumumab

will be summarized in all the tables specified in this section.

7.6.4 Adverse Events Leading to Dose Modification

AEs leading to dose delay/reduction will be summarized by treatment group:

- Overall summary of AEs leading to dose delay/reduction by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.
- Overall summary of related AEs leading to dose delay/reduction by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.

The analysis will be conducted using the 30-day safety window.

A by-subject AEs leading to dose delay/reduction listing will be provided.

7.6.5 Adverse Events

Adverse events will be summarized by treatment group.

The following analyses will be conducted using the 30 days safety window only:

- Overall summary of any AEs by worst CTC grade (1, 2, 3, 4, 5, not reported, total) presented by SOC/PT.

- Overall summary of any AEs presented by worst CTC grade (any grade, grade 3-4, grade 5) by SOC/PT. This table will be restricted to events with an incidence greater or equal to 5% in any treatment group.
- Overall summary of any non-serious AEs presented by SOC/PT. This table will be restricted to events with an incidence greater or equal to 5% in any treatment group.
- Overall summary of any AEs that required immune modulating medication by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.
- Overall summary of drug-related AEs by worst CTC grade (1, 2, 3, 4, 5, not reported, total) presented by SOC/PT.

The following analyses will be conducted using the 30 days safety window and repeated using the 100 days safety window:

- Overall summary of drug-related AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.

A by-subject AE listing will be provided. A by-subject listing of any AE requiring immune modulating medications will also be provided.

For those subjects treated in Arm B and for those intermediate/poor risk treated subjects in Arm A and Arm C:

The following analyses will be conducted using the 30 days safety window only:

- Overall summary of any AEs by worst CTC grade (1, 2, 3, 4, 5, not reported, total) presented by SOC/PT.
- Overall summary of any AEs that required immune modulating medication by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.
- Overall summary of drug-related AEs by worst CTC grade (1, 2, 3, 4, 5, not reported, total) presented by SOC/PT.

The following analyses will be conducted using the 30 days safety window and repeated using the 100 days safety window:

- Overall summary of drug-related AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.

A by-subject AE listing will be provided. A by-subject listing of any AE requiring immune modulating medications will also be provided

7.6.6 Select Adverse Events (EU/ROW Submissions)

Unless otherwise specified, analyses will be performed by select AE category. Analyses will also be repeated by subcategory of endocrine events.

7.6.6.1 Incidence of Select AE

Select AEs will be summarized by treatment group for each category/subcategory.

The following analyses will be conducted using the 30-day safety window only:

- Overall summaries of any select AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category or Subcategory/PT.
- Overall summaries of any drug-related select AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category or Subcategory/PT.
- Overall summaries of any serious select AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category or Subcategory /PT.
- Overall summaries of drug-related serious select AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category or Subcategory /PT.
- Overall summaries of any select AEs leading to discontinuation by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category or Subcategory /PT.
- Overall summaries of drug-related select AEs leading to discontinuation by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category or Subcategory /PT.
- Summary of frequency of unique select AEs by Category.

A by-subject select AE listing will be provided.

7.6.6.2 Time-to Onset of Select AE

Time-to onset of drug-related select AEs (any grade, grade 3-5) will be summarized for each category/subcategory by treatment group.

Time-to onset analyses are restricted to treated subjects who experienced at least one drug-related select AE in the category/subcategory. The analyses will be conducted using the 30-day safety window.

Additional details regarding the time-to onset definition are described in time-to onset definition subsection of APPENDIX 1.

7.6.6.3 Time-to Resolution of Select AE

Time-to resolution of the following specific events will be summarized separately for each category/subcategory.

- Time-to resolution of drug-related select AE (any grade, grade 3-5) by treatment group
- Time-to resolution of drug-related select AE (any grade, grade 3-5) where immune modulating medication was initiated, by treatment group

Time-to resolution analyses are restricted to treated subjects who experienced the specific events. Time-to resolution where immune modulating medication was initiated analyses are restricted to treated subjects who experienced the specific events and who received immune modulating medication during the longest select AE.

The analyses will be conducted using the 30-day safety window.

The following summary statistics will be reported: percentage of subjects with resolution of the longest select AE, median time-to resolution along with 95% CI (derived from Kaplan-Meier estimation) and ranges.

See time-to resolution definition subsection of APPENDIX 1 for additional details.

7.6.7 Immune-Mediated Adverse Events (US Submission)

IMAEs will be summarized by treatment group for each immune-mediated category / PT using the 100-day safety window:

- Overall summary of non-endocrine IMAEs by worst CTC grade (any grade, grade 3-4, grade 5) where immune modulating medication was initiated presented by Category / PT.
- Overall summary of endocrine IMAEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category / PT.
- Overall summary of non-endocrine IMAEs leading to discontinuation by worst CTC grade (any grade, grade 3-4, grade 5) where immune modulating medication was initiated presented by Category / PT.
- Overall summary of endocrine IMAEs leading to discontinuation by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category / PT.
- Overall summary of non-endocrine IMAEs leading to dose delay or reduction by worst CTC grade (any grade, grade 3-4, grade 5) where immune modulating medication was initiated presented by Category / PT
- Overall summary of endocrine IMAEs leading to dose delay or reduction by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category / PT.
- Summaries of time-to onset and time-to resolution of non-endocrine IMAEs where immune modulating medication was initiated presented by Category.
- Summaries of time-to onset and time-to resolution of endocrine IMAEs presented by Category.

A by-subject listing of IMAEs will be provided. By-subject listings of time-to resolution for longest IMAEs cluster (any grade and grade 3-5 in separate summaries) will also be provided. For new studies which collect investigator assessment of potential IMAE data, a by-subject listing of AEs considered as immune-mediated events per investigator but not qualified for IMAEs definition will also be provided.

In addition, for all nivolumab treated subjects who experienced at least one IMAE, the following data presentation will be provided:

- Summary of subjects who were re-challenged with nivolumab by IMAE category, with extended follow-up
- Summary of subjects who were re-challenged with nivolumab or ipilimumab by IMAE category, with extended follow-up

For these, re-challenge is considered to have occurred when last nivolumab and/or ipilimumab infusion was administered after the onset of an IMAE.

7.6.8 Other Events of Special Interest

OEOSI will be summarized by treatment group for each category.

The following analyses will be conducted using the 100-day safety window:

- Overall summary of OEOSI by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category / PT
- Overall summary of drug-related OEOSI by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category / PT

A by-subject listing of OEOSI will be provided.

7.6.9 Multiple Events

The following summary tables will be provided:

- A table showing the total number and rate (exposure adjusted) of occurrences for all AEs.
- A table showing the total number and rate (exposure adjusted) of occurrences for AEs occurring in at least 5% of subjects in any treatment group.

In addition, the rate (exposure adjusted) and its 95% CI evaluated for different time intervals will be displayed graphically for each treatment group. This analysis will be limited to the rate of all AEs and all drug-related AEs. The analyses will be conducted using the 30-day safety window.

A listing displaying the unique instances of all AEs, i.e., after duplicates have been eliminated and overlapping and contiguous occurrences of the same event (i.e. same PT) have been collapsed will be provided. No formal comparisons will be made between treatment groups.

7.6.10 Laboratory Parameters

The analysis population for each laboratory test is restricted to treated subjects who underwent that laboratory test. Laboratory tests (in addition to the tests specified below) with CTC criteria collected in the specific studies may also be included in the summaries.

A by-subject listing of differences in categorization of SI and US laboratory test results will be provided.

7.6.10.1 Hematology

The following will be summarized by treatment group as worst CTC grade on-treatment per subject and as shift table of worst on-treatment CTC grade compared to baseline CTC grade per subject: hemoglobin (HB), platelets, white blood counts (WBC), absolute neutrophils count (ANC) and lymphocyte count (LYMPH).

The analyses will be conducted using the 30-day safety window.

A by-subject listing of these laboratory parameters will be provided.

7.6.10.2 Serum Chemistry

The following will be summarized by treatment group as worst CTC grade on-treatment per subject and as shift table of worst on-treatment CTC grade compared to baseline CTC grade per subject: ALT, AST, alkaline phosphatase (ALP), total bilirubin and creatinine.

The analyses will be conducted using the 30-day safety window.

A by-subject listing of these laboratory parameters will be provided.

7.6.10.3 Electrolytes

The following will be summarized by treatment group as worst CTC grade on-treatment per subject and as shift table of worst on-treatment CTC grade compared to baseline CTC grade per subject: sodium (high and low), potassium (high and low), calcium (high and low), magnesium (high and low), and Glucose Serum (fasting hyperglycemia and hypoglycemia regardless of fasting status).

The analyses will be conducted using the 30-day safety window.

A by-subject listing of these laboratory parameters will be provided.

7.6.10.4 Additional Analyses

In addition, further analyses on specific laboratory parameters will be performed by treatment group:

Abnormal Hepatic Function Test

The number of subjects with the following laboratory abnormalities from on-treatment evaluations will be summarized by treatment group:

- ALT or AST > 3 x ULN, > 5 x ULN, > 10 x ULN and > 20 x ULN
- Total bilirubin > 2 x ULN
- ALP > 1.5 x ULN
- Concurrent (within 1 day) ALT or AST > 3 x ULN and total bilirubin > 1.5 x ULN
- Concurrent (within 30 days) ALT or AST > 3 x ULN and total bilirubin > 1.5 x ULN
- Concurrent (within 1 day) ALT or AST > 3 x ULN and total bilirubin > 2 x ULN
- Concurrent (within 30 days) ALT or AST > 3 x ULN and total bilirubin > 2 x ULN

The analyses will be conducted using the 30-day safety window.

A by-subject listing of these specific abnormalities will be provided.

Abnormal Thyroid Function Test

The number of subjects with the following laboratory abnormalities from on-treatment evaluations will be summarized by treatment group:

- TSH value > ULN and

  - with baseline TSH value ≤ ULN
  - with at least one FT3/FT4 test value < LLN within 2-week window after the abnormal TSH test
  - with all FT3/FT4 test values ≥ LLN within 2-week window after the abnormal TSH test
  - with FT3/FT4 missing within 2-week window after the abnormal TSH test.

- TSH < LLN and
  - with baseline TSH value ≥ LLN
  - with at least one FT3/FT4 test value > ULN within 2-week window after the abnormal TSH test
  - with all FT3/FT4 test values ≤ ULN within 2-week window after the abnormal TSH test
  - with FT3/FT4 missing within 2-week window after the abnormal TSH test

The analyses will be conducted using the 30-day safety window.

A by-subject listing of these specific abnormalities will be provided.

7.6.11 Vital Signs and Pulse Oximetry

Vital signs and pulse oximetry (i.e. % oxygen saturation) collected on the CRF will be provided in separate listings.

7.6.12 Physical Measurements

Physical measurements will be listed by subject.

7.6.13 Non-Protocol Medical Procedures

Non-protocol medical procedures will be listed by subject.

7.6.14 Immunogenicity Analysis

Further details on immunogenicity background and rationale, definitions, population for analyses and endpoints are described in APPENDIX 3.

Incidence of ADA

Number (%) of subjects will be reported for the following parameters based on Evaluable Subjects.

- Baseline ADA-positive
- ADA-positive
  - Persistent Positive (PP)
  - Not PP-Last Sample Positive
  - Other positive
  - Neutralizing Positive
- ADA-negative

A listing of all ADA assessments will be provided. A separate listing of ADA assessments for subjects with neutralizing positive will also be provided.

A spider plot of nivolumab ADA test result (titers) over time may be provided for nivolumab ADA positive subjects.

Clinical implications

Clinical implications of nivolumab immunogenicity will be primarily focused on subjects with persistent ADA-positive relative to ADA-negative. Subjects with any ADA-positive samples after initiation of treatment (relative to baseline) may be used to explore clinical implications. Effect of immunogenicity on clearance of nivolumab will be explored by comparison of clearance estimates (determined by PPK analysis). Effect of immunogenicity on safety will be explored by examining the frequency and type of AEs of special interest such as hypersensitivity/infusion reaction. Summary tables for incidence of overall and each of the preferred terms will be provided, if the number of subjects is of sufficient size (e.g., at least 10 subjects). Otherwise, individual subject's safety profile will be examined and described based on a listing. Clinical implications on efficacy will also be explored similarly. Association between trough concentrations of nivolumab or combination drug (e.g., ipilimumab) and ADA assessments may be explored, as needed.

The following data presentation will be provided:

- Swimmer plot of occurrence of ADA and NAb Occurrence in Relation to PFS, BOR per investigator and OS

7.6.15 Pregnancy

A by-subject listing of pregnancy tests results will be provided for randomized female subjects.

7.6.16 Adverse Events By Subgroup

Overall summary of any AEs and drug-related AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT and for each treatment group for the following subgroups:

- Sex (Male vs. Female)
- Race (White, Black,or African American, Asian, Other
- Age (< 65 vs. 65 - < 75 vs. 75 - < 85 vs. ≥ 85 vs. ≥ 75 vs. ≥ 65)
- Region (US/Canada/W.Europe/N.Europe vs. ROW)

These analyses will be conducted using the 30-day safety window only.

7.7 Pharmacokinetics

The nivolumab and cabozantinib concentration data obtained in this study will be combined with data from other studies in the clinical development program to develop a population PK model. This model will be used to evaluate the effects of intrinsic and extrinsic covariates on the PK of nivolumab and cabozantinib. In addition, exposure-response analyses with selected efficacy and

safety endpoints will be conducted. Results of population PK and exposure response-analyses will be reported separately.

7.8 Biomarkers

Analyses for PD-L1 are described below.

7.8.1 Distribution of PD-L1 Expression

Descriptive statistics of PD-L1 expression:

- Listing of all PD-L1 IHC data, all randomized subjects
- Summary of tumor specimen acquisition and characteristics, all randomized subjects
- Summary statistics of PD-L1 expression in all randomized subjects with quantifiable PD-L1 expression
- Summary of BOR and ORR by baseline PD-L1 expression at 1, 5, and 10% cutoffs in all randomized subjects
- Kaplan-Meier curves for PFS and OS by BICR by baseline expression of PD-L1 < 1%, >= 1 and < 5, > = 5 and < 10, and >= 10
- Cumulative distribution plot of baseline PD-L1 expression versus population percentile in all randomized subjects with quantifiable PD-L1 expression
- Box plots of PD-L1 expression at 1, 5, and 10% cutoffs versus Response Status (BICR assessment) in all randomized subjects with quantifiable PD-L1 expression
- Waterfall plot of individual PD-L1 expression in all randomized subjects with quantifiable PD-L1 expression

7.8.2 Other Exploratory Biomarkers

The following analyses for other exploratory biomarkers will be provided when the data becomes available, which may occur after the CSR is written and therefore a seperate Data Presentation Plan (DPP) will be developed:

Descriptive statistics of c-MET expression:

- Listing of all c-MET IHC data, all randomized subjects
- Summary of tumor specimen acquisition and characteristics, all randomized subjects.
- Summary statistics of c-MET expression in all randomized subjects with quantifiable c-MET expression
- Correlation statistics of BOR and ORR by baseline c-MET expression levels
- Cumulative distribution plot of baseline c-MET expression versus population percentile in all randomized subjects with quantifiable c-MET expression
- Kaplan-Meier curves for PFS and OS by BICR by baseline expression of tertile c-MET expression

- Box plots of c-MET expression at tertile cutoffs versus Response Status (BICR assessment) in all randomized subjects with quantifiable c-MET expression
- Waterfall plot of individual c-MET expression in all randomized subjects with quantifiable c-MET expression

Descriptive statistics of MDSC expression at baseline:

- Listing of all MDSC data, all randomized subjects
- Summary of tumor specimen acquisition and characteristics, all randomized subjects
- Summary statistics of MDSC expression in all randomized subjects with quantifiable MDSC expression
- Correlation statistics of BOR and ORR by baseline MDSC expression levels
- Cumulative distribution plot of baseline MDSC expression versus population percentile in all randomized subjects with quantifiable MDSC expression
- Kaplan-Myer curves for PFS and OS by BICR by baseline expression of tertitle MDSC expression
- Box plots of MDSC expression at tertile cutoffs versus Response Status (BICR assessment) in all randomized subjects with quantifiable MDSC expression
- Waterfall plot of individual MDSC expression in all randomized subjects with quantifiable MDSC expression

Descriptive statistics of serum cytokine expressions (only MIG, IP10 and IFN-g) changes in the course of treatment:

- Line plot per arm for each cytokine on X axis and with timeline on Y axis with statistical difference from the baseline to be marked

Descriptive statistics for gene expression signature analyses:

- Published, on-target and mechanism of action related pathway related gene expression signatures (e.g. CD8, angiogenesis etc) will be assessed in BOR, ORR, PFS and OS relationship

7.9 Clinical Outcomes Assessments

The analysis of FKSI-19 and EQ-5D-3L will be restricted to randomized subjects in Arm A and C who have an assessment at baseline and at least one post-baseline assessment.

7.9.1 FKSI-19

Unless otherwise specified, the analysis of FKSI-19 will be performed by treatment group in all randomized subjects who have an assessment at baseline and at least one or more post-baseline assessments. The following descriptive analyses will be conducted:

- Questionnaire completion rate, defined as the proportion of questionnaires actually received out of the expected number (i.e., number of subjects on treatment or in follow up), will be calculated and summarized for each assessment time point.
- For the total score and subscales, separately:
  - Mean score and mean change from baseline at each assessment time point will be summarized using descriptive statistics (N, mean, SD, median, 25th and 75th percentiles, minimum, maximum).
  - A plot summarizing the mean change from baseline will be presented and including 95% CI.

By subject listing of FKSI-19 will be provided.

7.9.2 EuroQol EQ-5D-3L

The following descriptive analyses will be conducted:

- EQ-5D-3L questionnaire completion rate, defined as the proportion of questionnaires actually received out of the expected number (i.e. number of subjects on treatment or in follow up), will be calculated and summarized for each assessment time point by treatment group.
- A by-subject listing of the level of problems in each dimension, corresponding to EQ-5D-3L health state (i.e., 5-digit vector), EQ-5D-3L utility index score, and EQ-5D-3L VAS score will be provided.
- Proportion of subjects reporting problems for the 5 EQ-5D-3L dimensions at each assessment time point will be summarized by level of problem and by treatment group. Percentages will be based on number of subjects assessed at assessment time point.
- For the EQ-5D-3L utility index and VAS scores, separately:
  - Mean score and mean change from baseline at each assessment time point will be summarized by treatment group using descriptive statistics (N, mean with SD and 95% CI, median, first and third quartiles, minimum, maximum).
  - A line graph summarizing the mean changes from baseline will be produced.

8 CONVENTIONS

The following conventions may be used for imputing partial dates for analyses requiring dates:

- For missing and partial adverse event onset dates, imputation will be performed using the Adverse Event Domain Requirements Specification[16]
- For missing and partial adverse event resolution dates, imputation will be performed as follows (these conventions may change):
  - If only the day of the month is missing, the last day of the month will be used to replace the missing day. If the imputed date is after the death date or the last known alive date, then the latest known alive date or death date is considered as the resolution date.

  - If the day and month are missing or a date is completely missing, it will be considered as missing.

- Missing and partial non-study medication domain dates will be imputed using the derivation algorithm described in 4.1.3 of BMS Non-Study Medication Domain Requirements Specification[17].

- Missing and partial radiotherapy and surgery dates will be imputed using algorithm described in APPENDIX 2.

- For death dates, the following conventions will be used for imputing partial dates:
  - If only the day of the month is missing, the 1st of the month will be used to replace the missing day. The imputed date will be compared to the last known alive date and the maximum will be considered as the death date.
  - If the month or the year is missing, the death date will be imputed as the last known alive date.
  - If the date is completely missing but the reason for death is present, the death date will be imputed as the last known date alive.

- For date of progression after start of study therapy, the following conventions will be used for imputing partial dates:
  - If only the day of the month is missing, the 1st of the month will be used to replace the missing day. In case of the date of death is present and complete, the imputed progression date will be compared to the date of death. The minimum of the imputed progression date and date of death will be considered as the date of progression.
  - If the day and month are missing or a date is completely missing, it will be considered as missing.

- For date of progression to prior therapies, the following conventions will be used for imputing partial dates:
  - If only the day of the month is missing, the 1st of the month will be used to replace the missing day.
  - If the day and month are missing or a date is completely missing, it will be considered as missing.

- For other partial/missing dates, the following conventions were used:
  - If only the day of the month is missing, the 15th of the month will be used to replace the missing day.
  - If both the day and the month are missing, “July 1” will be used to replace the missing information.
  - If a date is completely missing, it will be considered as missing.

The following conversion factors will be used to convert days to months or years:

1 month = 30.4375 days and 1 year = 365.25 days.

Duration (e.g. time-to onset, time-to resolution) will be calculated as follows:

Duration = (Last date - first date + 1)

Last known alive date will be defined based on all appropriate dates collected on the CRF.

All statistical analyses will be carried out using SAS (Statistical Analysis System software, SAS Institute, North Carolina, USA) unless otherwise noted.

9 CONTENT OF REPORTS

All analyses described in this SAP will be included in the Clinical Study Report(s) except where otherwise noted. Refer to the Data Presentation Plan for mock-ups of all tables and listings.

10 DOCUMENT HISTORY

Table 10-1: Document History

Version Number	Author(s)	Description
1	Burcin Simsek	Original issue

APPENDIX 1 TIME-TO ONSET AND TIME-TO RESOLUTION DEFINITION AND CONVENTIONS FOR SELECT ADVERSE EVENTS, IMMUNE-MEDIATED ADVERSE EVENTS AND EVENTS OF SPECIAL INTEREST

Time-to onset definition

Time-to onset of AE (any grade) for a specific category is defined as the time between the day of the first dose of study treatment and the onset date of the earliest AE (of any grade) in this category.

The time-to onset of AE (grade 3-5) for a specific category is defined similarly with an onset date corresponding to a grade 3-5 AE.

Time-to onset of drug-related AE (any grade or grade 3-5) for a specific category is defined similarly but restricted to drug-related AE.

Time-to onset for a specific subcategory is defined similarly but restricted to event of this subcategory.

Time-to resolution definition

In order to derive the time-to resolution, overlapping or contiguous AEs within a specific category or subcategory will be collapsed into what will be termed "clustered" AEs. For example, if a subject (without pre-treatment AE) experienced an AE from 1st to 5th January, another AE (with different PT but within same category) from 6th to 11th January and same AE from 10th to 12th January, these will be collapsed into one clustered AE from 1st to 12th January.

Key derivation steps for each type of clustered AEs is summarized as follows:

- For any grade AE: Collapse any on-treatment AE from the same category
- For drug-related any grade AE: Collapse any on-treatment drug-related AE from the same category
- For grade 3-5 AE: Collapse any on-treatment AE from the same category. Resolution will be based on the onset date of the earliest grade 3-5 records (if no grade 3-5 record, clustered AE is excluded)
- For drug-related grade 3-5 AE: Collapse any on-treatment drug-related AE from the same category. Resolution will be based on the onset date of the earliest grade 3-5 records (if no grade 3-5 record, clustered AE is excluded)

Time-to resolution of AE (any grade) for a specific category is defined as the longest time from onset to complete resolution or improvement to the grade at baseline among all clustered AEs experienced by the subject in this category per adverse event criteria category. Events which worsened into grade 5 events (death) or have a resolution date equal to the date of death are considered unresolved. If a clustered AE is considered as unresolved, the resolution date will be

censored to the last known alive date. Improvement to the grade at baseline implies that all different events in the clustered adverse event should at least have improved to the corresponding (i.e. with same preferred term) baseline grade. This measure is defined only for subjects who experienced at least one AE in the specific category.

The time-to resolution of AE (grade 3-5) for a specific category is defined similarly with an onset date corresponding to a grade 3-5 AE.

Time-to resolution of drug-related AE (any grade or grade 3-5) for a specific category is defined similarly but restricted to drug-related AE.

The time-to resolution of AE (any grade or grade 3-5, drug-related or all) where immune modulating medication was initiated is defined similarly. For data presentation not restricted to IMAE, the additional condition that the subject started an immune modulating medication during the longest AE resolution period will be applied.

Time-to resolution for a specific subcategory is defined similarly but restricted to event of this subcategory.

The algorithm for collapsing adverse event records is using the following conventions:

For each subject and specified category, the corresponding adverse event records will be collapsed when:

1) Multiple adverse event records have the same onset date.
2) The onset date of an event record is either the same day or 1 day later than the resolution date of a preceding event record (contiguous events).
3) The onset date of an event record is after the onset date and prior to or on the resolution date of a preceding event record (overlapping events).

APPENDIX 2 MISSING AND PARTIAL RADIOTHERAPY AND SURGERY DATES IMPUTATION ALGORITHMS

Procedures – Imputation Rules.

If reported procedure start date is a full valid date then set start date equal to the date part of procedure start date.

In case of partial date use imputation rules described below:

- If only day is missing then
  - If month and year of procedure match month and year of first dose date then impute as date of first dose;
  - If month and year of procedure don't match month and year of first dose date then impute as first day of that month and year.
- If both day and month are missing, then impute as maximum between 01JAN of the year and date of the first dose;
- If date is completely missing or invalid then leave missing.

Note: Imputation is not applicable to data where start date is not collected (for example "PRIOR RADIOTHERAPY" CRF). Set start date to missing in this case.

If reported end date is a full valid date then set end date equal to the date part of the reported end date.

In case of partial date use imputation rules described below:

- If reported end date is partial then set end date equal to the last possible reported end date based on the partial entered reported end date.
- If reported end date is missing, continuing, unknown or invalid then set end date equal to the most recent database extraction date.

If end date was imputed then compare end date to the death date or last known alive date if subject is not dead. If posterior then end date should be imputed to death date (or last known alive date if subject not dead).

Note: Imputation of partial dates only applies to data entered on "RADIOTHERAPY" CRF page. For other CRF pages in case of partial dates set end date to missing.

Surgeries – Imputation Rules.

If reported surgery date is a full valid date then set start date equal to the date part of surgery date.

In case of partial date, use one of the two imputation rules described below:

A. For data collected on "PRIOR SURGERY RELATED TO CANCER" CRF page:

- If only day is missing then impute as the first day of the month;
- If both day and month are missing then then impute as 01JAN of the year;

- If date is completely missing or invalid then leave missing.

B. For data collected on other CRF pages (deemed to be on-treatment/subsequent surgeries):

- If only day is missing then
  - If month and year of surgery match month and year of first dose date then impute the missing date as the date of first dose;
  - If month and year of surgery don't match month and year of first dose date then impute as first day of that month and year;
- If both day and month are missing then impute as maximum between 01JAN of the year and date of the first dose;
- If date is completely missing or invalid then leave missing.

APPENDIX 3 IMMUNOGENICITY ANALYSIS: BACKGROUND AND RATIONALE

The following summary is from the FDA Guidance for Industry Immunogenicity Assessment for Therapeutic Protein Products and White Paper on Assessment and Reporting of the Clinical Immunogenicity of Therapeutic Proteins and Peptides – Harmonized Terminology and Tactical Recommendations by Shankar et al. The program-level definitions of sample- and subject-level ADA status are based on recommendation from the BMS Immunogenicity Council.

Immune responses to therapeutic protein products may pose problems for both subject safety and product efficacy. Immunologically based adverse events, such as anaphylaxis and infusion reactions, have caused termination of the development of therapeutic protein products or limited the use of otherwise effective therapies. Unwanted immune responses to therapeutic proteins may also neutralize the biological activity of therapeutic proteins and may result in adverse events not only by inhibiting the efficacy of the therapeutic protein product, but by cross-reacting to an endogenous protein counterpart, if present. Because most of the adverse effects resulting from elicitation of an immune response to a therapeutic protein product appear to be mediated by humoral mechanisms, circulating antibody has been the chief criterion for defining an immune response to this class of products.

ADA is defined as biologic drug-reactive antibody, including pre-existing host antibodies that are cross-reactive with the administered biologic drug (baseline ADA). Titer is a quasiquantitative expression of the level of ADA in a sample. By employing a serial dilution-based test method, titer is defined as the reciprocal of the highest dilution of the sample (e.g., dilution of 1/100 = titer of 100). The ADA is also tested, via a cell-based biologic assay or a non cell-based competitive ligand-binding assay for a subpopulation of ADA known as neutralizing antibodies (NAb), which inhibits or reduces the pharmacological activity of the biologic drug molecule regardless of its in vivo clinical relevance. Non-neutralizing ADA (non-NAb) is ADA that binds to the biologic drug molecule but does not inhibits its pharmacological activity.

ADA should be tested using sensitive and valid methods and employing an appropriate strategy for elucidating immunogenicity. Detection of ADA is typically performed in three tiers (screening, confirmatory, and titer) using statistically determined cutpoints and samples testing positive in the ADA assay are analyzed for neutralizing activity, especially in late-stage clinical studies. “Detection” of ADA implies that drug-specific ADA was confirmed. The ‘drug tolerance” of an assay (highest drug concentration that does not interfere in the ADA detection method) is not an absolute value and differs between individuals due to the varying avidities of ADA immune responses. An ADA sampling strategy of collecting samples at times when the least drug concentration is anticipated (trough concentrations) can increase the likelihood of accurate ADA detection.

It is useful to present ADA results from clinical studies as (a) characteristics of the ADA immune response, (b) relationship of ADA with pharmacokinetics (PK) and, when relevant, pharmacodynamics (PD) biomarkers, and (c) relationship of ADA with clinical safety and efficacy.

Clinical consequences of ADA can range from no apparent clinical effect to lack of efficacy (primary treatment failure), loss of efficacy (secondary treatment failure) or heightened effect due to altered exposure to the biologic drug, adverse drug reactions (administration-related systemic or site reactions), and severe adverse drug reactions (anaphylaxis and unique clinical problems associated with cross-reactivity and neutralization of endogenous molecules). Thus it becomes important to examine any associations between ADA or any of its attributes with the various clinical sequelae. The presence of ADA may or may not preclude the administration of drug to ADA-positive subjects because the outcome is dependent upon the magnitude of the impact of ADA on PK and PD. Hence, the relationship of ADA with PK/PD is an important additional consideration, but does not necessarily result in a clinically impactful consequence per se.

Immunogenicity Endpoints

A fundamental metric that informs clinical immunogenicity interpretation is the incidence of ADA in a study or across comparable studies. ADA incidence is defined as the proportion of the study population found to have seroconverted or boosted their pre-existing ADA during the study period.

Terms and Definitions

Validated ADA test methods enable characterization of samples into ADA-positive vs. ADA-negative. To classify the ADA status of a subject using data from an in vitro test method, each sample from the subject is categorized based on the following definitions:

Sample ADA Status**:**

- Baseline ADA-positive sample: ADA is detected in the last sample before initiation of treatment
- Baseline ADA-negative sample: ADA is not detected in the last sample before initiation of treatment
- ADA-positive sample: After initiation of treatment, (1) an ADA detected (positive seroconversion) sample in a subject for whom ADA is not detected at baseline, or (2) an ADA detected sample with ADA titer to be at least 4-fold or greater (≥) than baseline positive titer
- ADA-negative sample: After initiation of treatment, ADA not positive sample relative to baseline

Next, using the sample ADA status, subject ADA status is defined as follows:

Subject ADA Status:

- Baseline ADA-positive subject: A subject with baseline ADA-positive sample
- **ADA-positive subject**: A subject with at least one ADA positive-sample relative to baseline at any time after initiation of treatment
  1) *Persistent Positive (PP):* ADA-positive sample at 2 or more consecutive time points, where the first and last ADA-positive samples are at least 16 weeks apart
  2) *Not PP-Last Sample Positive:* Not persistent positive with ADA-positive sample at the last sampling time point

3) *Other Positive:* Not persistent positive but some ADA-positive samples with the last sample being negative
4) *Neutralizing Positive:* At least one ADA-positive sample with neutralizing antibodies detected

- **ADA-negative subject:** A subject with no ADA-positive sample after the initiation of treatment.

(Note: 16 weeks was chosen based on a long half-life of IgG4.)

Population for Analyses

Analysis of immunogenicity data will be based on ADA evaluable subjects defined as all treated subjects with baseline and at least 1 post-baseline immunogenicity assessment. Analysis dataset and data listing will include all available ADA samples. However, subject-level ADA status will be defined based on only adequate samples (e.g., excluding 1-hour post-infusion samples when clearly indicated).

11 REFERENCES

[1] Apolo AB, Tomita Y, Lee MJ, Lee S, Frosch A, Steinberg SM. Effect of Cabozantinib on immunosuppressive subsets in metastatic urotherlial carcinoma [abstract No. 4501]. 2014. J Clin Oncol. 2014; 32 (5s). 2014 ASCO Annual Meeting.

[2] Dolan P. Modeling valuations for EuroQol health states. Medical Care 1997;35: 1095-1108

[3] The EuroQol Group: EuroQol: A new facility for the measurement of health-related quality of life—The EuroQol Group. Health Policy 16:199-208, 1990.

[4] Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. Health Qual Life Outcomes 2007;5:70.

[5] Brookmeyer R. and Crowley J. A confidence interval for the median survival time. Biometrics 38:29-41, 1982

[6] Klein, J. P. and Moeschberger, M. L. (1997), Survival Analysis: Techniques for Censored and Truncated Data, New York: Springer-Verlag

[7] Clopper, CJ and Pearson, ES. The use of confidence or fiducial limits illustrated in the case of the binomial.Biometrika 26: 404-423, 1934.

[8] Fagerland MW, Lydersen S, Laake P. Recommended confidence intervals for two independent binomial proportions. Stat Methods Med Res 2011. Published Online First 13 Oct 2011.

[9] Global Biometric Sciences, SAS Analysis Dataset Specification, Unique Adverse Events. Version 2.0, April 22, 2015

[10] Glimm E, Maurer W, Bretz F. Hierarchical testing of multiple endpoints in group-sequential trials. Statistics in Medicine. 2010;29:219-228

[11] Greenwood, M. The errors of sampling of the survivorship tables, Reports on Public Health and Statistical Subjects, 33, Appendix 1, HMSO, London, 1926

[12] Kalbfleisch, J. D. and Prentice, R. L. (1980), The Statistical Analysis of Failure Time Data, New York: John Wiley & Sons.

[13] Fleming TR, Harrington DP 1991: Counting processes and survival analysis. New York: John Wiley and Sons

[14] Gail, M. and Simon, R. “Testing for qualitative interactions between treatment effects and subject subsets”. Biometrics, 1985; 41:361-372.

[15] Clopper, C.; Pearson, E. S. (1934). "The use of confidence or fiducial limits illustrated in the case of the binomial". *Biometrika* **26**: 404–413.

[16] Adverse Event Domain Requirements Specification Bristol Myers Squibb Co. PRI. Version 2.3.0 April 23, 2018

[17] Non-Study Medication Domain Requirements Specification Bristol Myers Squibb Co. PRI. Version 2.10.0 April 23, 2018

STATISTICAL ANALYSIS PLAN

A PHASE 3, RANDOMIZED, OPEN-LABEL STUDY OF NIVOLUMAB COMBINED WITH CABOZANTINIB VERSUS SUNITINIB IN PARTICIPANTS WITH PREVIOUSLY UNTREATED ADVANCED OR METASTATIC RENAL CELL CARCINOMA

PROTOCOL CA2099ER

VERSION # 2.0

DATE: 11-DEC-2019

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 BACKGROUND AND RATIONALE

Although multiple agents are approved as monotherapies for the treatment of subjects with metastatic renal cell carcinoma (mRCC), the testing of combination therapies, in particular, treatment with immune-checkpoint inhibitors in combination with tyrosine kinase inhibitors (TKIs) has not been fully explored. While single agent therapies have improved outcomes, ongoing drug resistance and disease progression demonstrate an urgent need to find more effective therapies for mRCC subjects.. The CA2099ER trial will include previously untreated subjects with advanced RCC and uses the well-characterized immune checkpoint inhibitor nivolumab in combination with cabozantinib, a known standard-of-care in mRCC subjects. Nivolumab combined with cabozantinib may be an important step forward in evaluating combination regimens which could potentially optimize the management of previously untreated subjects with mRCC.

Cabozantinib was shown to improve PFS and OS compared to everolimus, leading to its regulatory approval in a randomized phase 3 trial in subjects with advanced RCC that had progressed after anti-VEGFR therapy. Subsequently, a randomized phase 2 trial of cabozantinib vs sunitinib has demonstrated an improvement in ORR and PFS in intermediate- and poor-risk subjects with previously untreated mRCC (see Section 3.2.1.4 of the protocol). Cabozantinib has also been demonstrated to have effects on immune cells. In a study of 24 subjects with advanced urothelial carcinoma, cabozantinib treatment resulted in a decrease in circulating Tregs and increased PD-1 expression on Tregs. Low Tregs at baseline were also predictive of improved response to cabozantinib and survival.[1]

Given the promising clinical activity of cabozantinib in previously untreated mRCC and its potential immune effects, combining cabozantinib with nivolumab (in a doublet regimen) is a rational strategy to optimize first-line therapy in mRCC. An ongoing phase 1 study is evaluating both the doublet and triplet (nivolumab and ipilimumab combined with cabozantinib) regimens in subjects with refractory advanced urothelial cancer or other genitourinary tumors, including mRCC, and has defined dosing for both regimens that produces acceptable safety and tolerability (Section 3.2.1.5 of the protocol). CA209-9ER, 2-arm randomized phase 3 trial, will determine if the combination doublet regimen (nivolumab combined with cabozantinib) produces greater clinical benefit than sunitinib, a standard of care agent for subjects with previously untreated mRCC. In addition, this trial will reveal the adverse event profiles, quality of life measures, as well as exploratory biomarkers associated with these different first-line treatment regimens.

Research Hypothesis:

Treatment with nivolumab combined with cabozantinib (doublet regimen) will demonstrate an improvement in PFS per BICR compared to sunitinib monotherapy in subjects with previously untreated mRCC.

Schedule of Analyses:

This study will be monitored by an independent Data Monitoring Committee (DMC). Details are specified in the DMC charter.

The PFS analysis will occur after approximately 9-10 months minimum follow-up on all randomized subjects, which will be triggered by approximately 350 events from Arm A (nivolumab combined with cabozantinib; doublet regimen) and Arm C (sunitinib).

Two interim analyses of OS are planned. The first interim analysis is planned at the time of final PFS analysis and expected after observing 165 deaths among the randomized subjects in Arm A and Arm C (65% of the targeted OS events for final analysis). The second interim analysis is expected after observing 211 deaths among the randomized subjects in Arm A and Arm C (83% of targeted OS events needed for final analysis. The final analysis of OS is expected after observing 211 deaths among the randomized subjects in Arm A and Arm C.

The final PFS analysis will not occur prior to these conditions being met:

- at least 8 months minimum follow-up on all randomized subjects;
- at least 283 PFS events, which provide at least 90% power to detect a HR of 0.68 for PFS of Arm A versus Arm C; and
- at least 149 OS events, which provide 66% power if the observed HR for OS was 0.60.

The expected PFS analysis will occur at approximately 29 months from FPFV. The second interim and final analyses of OS are expected to occur approximately 34 months and 40 months from FPFV.

Secondary endpoints (including both efficacy endpoints OS and ORR) will be analyzed at the time of the final analysis of PFS based on a hierarchical testing strategy. In the event that the interim analysis for superiority of overall survival is positive, final (CSR) analyses will be performed prior to achieving 254 deaths; additional details can be found in section 7.5.7.

2 STUDY DESCRIPTION

Implementation of CA2099ER Global Revised Protocol 01 stops further randomization into Arm B (nivolumab + ipilimumab combined with cabozantinib). Subjects previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinical planned events, per protocol.

This is an open label, randomized trial of nivolumab combined with cabozantinib (doublet regimen) versus sunitinib in subjects with previously untreated (first line) advanced or metastatic RCC. Subjects will be randomized between Arm A and Arm C in a 1:1 ratio with approximately 638 subjects (319 per arm) capped at approximate 25% to represent the normal frequency of favorable risk group in mRCC. The rest of the randomized subjects will provide approximately 478 intermediate/poor risk randomized subjects (239 per arm). Subjects will be stratified at the time of randomization by IMDC prognostic score (0 [favorable risk] versus 1-2 [intermediate risk] versus 3-6 [poor risk]), PD-L1 tumor expression (≥ 1% versus < 1% or indeterminate), and region (US/Canada/Western Europe/Northern Europe versus rest of the world [ROW]).

The subject is randomly assigned to 1 of the 2 treatment arms as noted in the study schematic below.

- Arm A (Doublet): Nivolumab 240 mg flat dose IV Q2W + Cabozantinib 40 mg PO QD
  - Nivolumab to be continued until disease progression, unacceptable toxicity, or a maximum treatment of 2 years from the first dose in Cycle 1.
  - Cabozantinib to be continued until disease progression or unacceptable toxicity.
- Arm C: Sunitinib 50 mg PO QD for 4 weeks, followed by 2 weeks off, per cycle. Cycles to be continued until progression or unacceptable toxicity.

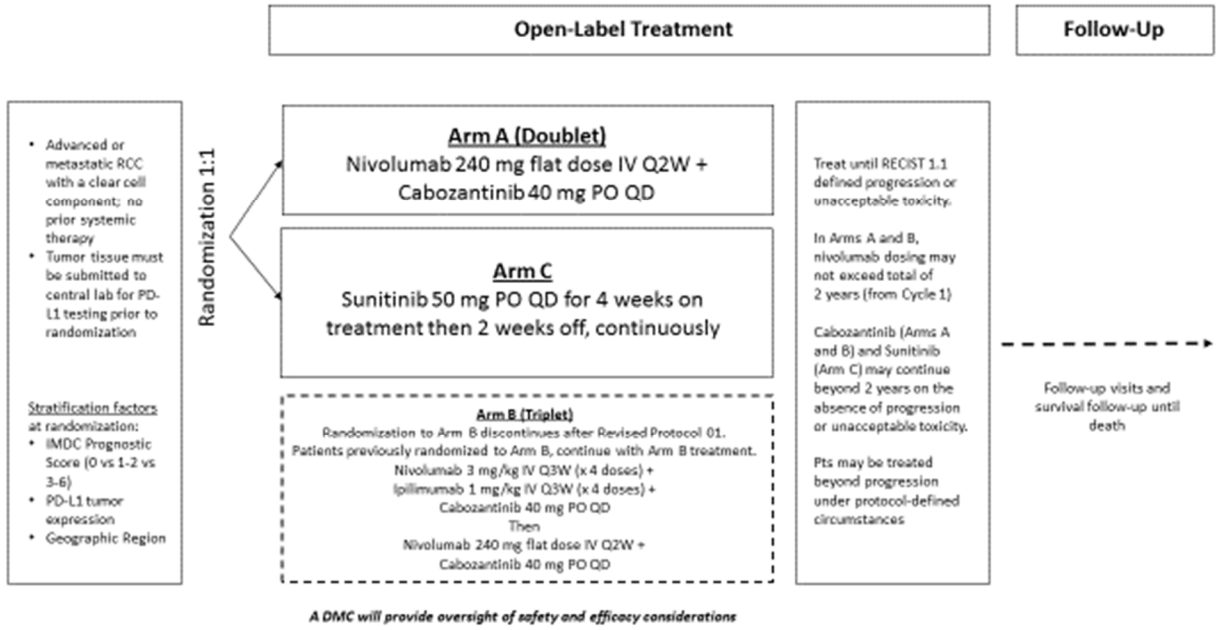
  Note - Randomization to Arm B stops with implementation of approved CA2099ER Global Revised Protocol 01. Treatment B continues only for subjects randomized to Arm B prior to implementation of Global Revised Protocol 01.

  - Arm B (Triplet): Nivolumab 3mg/kg IV + Ipilimumab 1 mg/kg IV, both Q3W x 4 doses + Cabozantinib 40 mg PO QD
    - Then Nivolumab 240 mg flat dose IV Q2W + Cabozantinib 40 mg PO QD.
    - Nivolumab to be continued until disease progression, unacceptable toxicity, or a maximum of 2 years from the first dose in Cycle 1.
    - Cabozantinib to be continued until progression or unacceptable toxicity.

Once randomized subjects in Arm A will continue nivolumab until progression, unacceptable toxicity, withdrawal of consent, or a maximum of 2 years from the first dose in Cycle 1, whichever occurs first. Cabozantinib (Arm A) may be continued until progression, unacceptable toxicity, or withdrawal of consent, whichever occurs first, and may extend beyond 2 years from the first dose in Cycle 1.

The study design schematic is presented in the figure below.

Figure 2-1: Study Design Schematic



2.1 Treatment Assignment

After the subject's initial eligibility is established and informed consent has been obtained, the subject must be enrolled into the study by accessing an Interactive Response Technologies web-based system (IRT) to obtain the subject number. All subjects will be centrally randomized using an Interactive Response Technology (IRT). Before the study is initiated, each user will receive log in information and directions on how to access the IRT.

Every subject that signs the informed consent form must be assigned a subject number in IRT. The investigator or designee will register the subject for enrollment by following the enrollment procedures established by BMS. The following information is required for enrollment:

- Date that informed consent was obtained
- Date of birth
- Sex at birth.

Once enrolled in IRT, enrolled subjects that have met all eligibility criteria will be ready to be randomized through the IRT. The following information is required for subject randomization:

- Subject number
- Date of birth
- IMDC Prognostic Score (0 versus 1-2 versus 3-6)
- Region (US/Canada/W Europe/N Europe versus ROW)

- PD-L1 tumor expression (≥ 1% versus < 1% or indeterminate)

Subjects meeting all eligibility criteria will be randomized in a 1:1 ratio to Arm A (nivolumab combined with cabozantinib) or Arm C (sunitinib), stratified by the following factors:

- IMDC Prognostic Score: 0 versus 1-2 versus 3-6
- Region: US/Canada/W Europe/N Europe versus ROW
- PD-L1 tumor expression: ≥ 1% versus < 1% or indeterminate

The randomization procedures will be carried out via permuted blocks within each stratum.

2.2 Blinding and Unblinding

This is an open label study.

2.3 Protocol Amendments

Table 2.3-1: Protocol Amendments

Document	Date of Issue	Summary of Change
Revised Protocol 02	06-MAY-2019	Major Changes: • Revised protocol 02 adjusts the timing of the PFS and OS interim analyses with modified hypothesized OS hazard ratio (HR). The number of randomized subjects is increased. • The interim analysis for ORR is removed, resulting in revised overall alpha for PFS and OS endpoints. • No change in eligibility or study procedure. • Clinical data for nivolumuab + ipilimumab in RCC has been updated. Other changes include more detail on PRO measures and updates to align with BMS standards for the nivolumab program.
Revised Protocol 01	18-DEC-2017	Primary revisions: (i) To stop enrollment into Arm B (nivolumab, ipilimumab and cabozantinib triplet) and (ii) to include favorable risk subjects (capped at 25%) in the primary data analysis. Secondary items include: (i) to add a Data Monitoring Committee review after 30 subjects are treated for 6 weeks, (ii) to adjust, clarify and add exclusion criteria, (iii) to add treatment restrictions, (iv) to clarify criteria associated with hemorrhage with regard to resuming treatment, (v) to specify an additional precaution when sunitinib dosing is resumed, and (vi) to apply newly updated Sponsor standards for nivolumab clinical protocols. Tertiary items include (i) incorporation of Administrative Letter 01 and (ii) correction of typographical and grammatical errors.

2.4 Data Monitoring Committee

An independent Data Monitoring Committee (DMC) has been established to provide oversight of safety and efficacy considerations, study conduct, and risk-benefit ratio. Following review, the DMC will recommend continuation, modification, or discontinuation of this study based on reported safety and efficacy data. Details of DMC responsibilities and procedures are specified in the DMC charter. Representatives of the Sponsor will serve only as coordinators of the committee, without having full member responsibilities or privileges. In addition, the Sponsor will independently review safety data in a blinded manner during the conduct of this trial to ensure that any safety issues are identified and addressed.

The DMC will conduct the first review of the safety data after at least 30 subjects are treated and followed for at least 6 weeks. The DMC will conduct its second review of the safety data after at least 75 subjects are treated and followed for at least 6 weeks. The DMC will conduct its review of the safety data focusing on the initial approximately 12 Japanese subjects (6 per arm) treated and followed for at least 4 weeks. The DMC will then review safety and the available efficacy data pertaining to primary endpoint to evaluate safety in the context of benefit, every six months thereafter.

The DMC will also review the formal final analysis of PFS (as per BICR) and first interim analysis of superiority of OS scheduled at around 29 months from FPFV. A second interim analysis of OS will be at around 34 months from FPFV. BMS will remain blinded to OS interim analyses unless DMC decides to disclose the formal interim analysis to BMS.

Details of the interim analyses can be found in section 7.5.7.

2.5 Blinded Independet Central Review

A blinded independent central review (BICR) committee has been established to provide an independent imaging review of images obtained in subjects participating in this study. Details of BICR responsibilities and processes may be found in the BICR Charter. The BICR determined PFS and ORR endpoints will be utilized as a part of primary and secondary efficacy analyses.

3 OBJECTIVES

3.1 Primary

- To compare PFS per BICR of nivolumab combined with cabozantinib (Arm A: doublet) with sunitinib (Arm C) in all randomized subjects.

3.2 Secondary

- To compare overall survival (OS) of Arm A with Arm C in all randomized subjects.
- To evaluate the objective response rate (ORR) of Arm A with Arm C per BICR in all randomized subjects.
- To assess overall safety and tolerability in all treated subjects.

3.3 Exporatory

- To explore potential predictive biomarkers of clinical response to nivolumab and cabozantinib combination.
- To evaluate health related quality of life (HRQoL).
- To characterize the pharmacokinetics of nivolumab and cabozantinib and explore exposure response relationships, if applicable.
- To characterize the immunogenicity of nivolumab.
- To assess PFS after next line of treatment (PFS2) in each arm.

4 ENDPOINTS

4.1 Primary Endpoints

Progression-free survival (PFS) is the primary endpoint. Two definitions are used for the analysis of PFS. The primary definition accounts for subsequent therapy by censoring at the last evaluable tumor assessment on or prior to the date of subsequent therapy. The secondary definition is irrespective of subsequent therapy and does not account for subsequent therapy.

Clinical deterioration in the absence of unequivocal evidence of progression (per RECIST v1.1 criteria) is not considered progression for purposes of determining PFS.

PFS rate at time *T* is defined as the probability that a subject has not progressed and is alive at time *T* following randomization. PFS rates at fixed time points (e.g. 6 months, depending on the minimum follow-up) are defined as the probability that a subject has not progressed and is alive at time *T* following randomization.

The first on-study tumor assessment is scheduled to be conducted at 12 weeks (± 1 week) following randomization. Subsequent tumor assessments are scheduled every 6 weeks (± 1 week) up to week 60, then every 12 weeks (± 2 weeks) until disease progression.

4.1.1 Primary Definition of Progression-Free Survival (Accounting for Subsequent Therapy)

The primary definition of PFS (PFS truncated at subsequent therapy, which includes anti-cancer therapy, tumor directed radiotherapy, or tumor directed surgery) is defined as the time between the date of randomization and the date of first documented tumor progression, based on BICR assessments (per RECIST v1.1 criteria), or death due to any cause, whichever occurs first.

Subjects who die without a reported progression will be considered to have progressed on the date of their death. The following censoring rules will be applied for the primary definition of PFS:

- Subjects who did not progress or die will be censored on the date of their last evaluable tumor assessment.
- Subjects who did not have any on study tumor assessments and did not die will be censored on their date of randomization.

- Subjects who receive subsequent anti-cancer therapy prior to documented progression will be censored at the date of the last evaluable tumor assessment conducted on or prior to the date of initiation of the subsequent anti-cancer therapy.
- Subjects who did not have a documented progression and received subsequent anti-cancer therapy will be censored at the date of the last evaluable tumor assessment conducted on or prior to the initiation of the subsequent anti-cancer therapy.

Censoring rules for the primary definition of PFS (PFS truncated at subsequent therapy) are presented as follows and in Table 4.1.1-1.

Figure 4.1.1-1: PFS Primary Definition

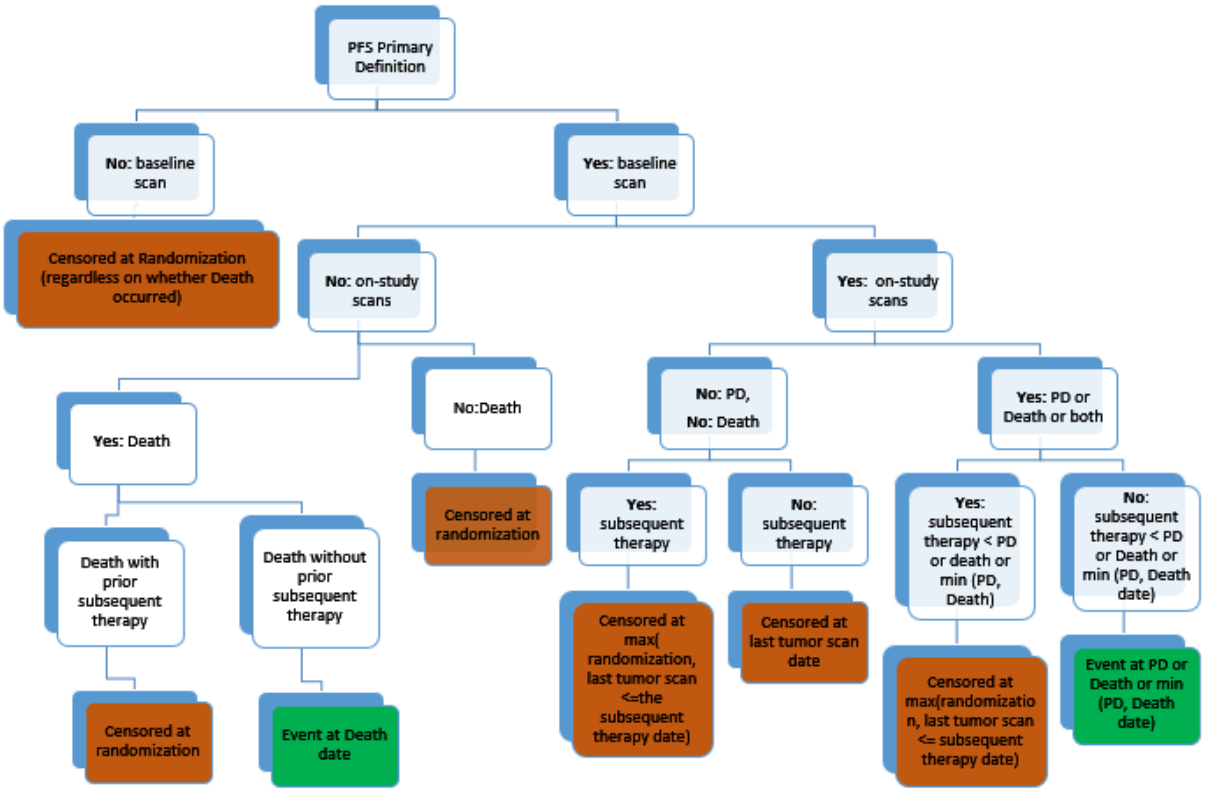


Table 4.1.1-1: Censoring Scheme Used in Primary Definition of PFS

Situation	Date of Progression or Censoring	Outcome
No baseline tumor assessments*	Date of randomization	Censored
No on study tumor assessments and no death*	Date of randomization	Censored
Subsequent anti-cancer therapy started without death or progression per RECIST v1.1 reported prior or on the same day	Date of last evaluable tumor assessment prior to or on the date of initiation of the subsequent anti-cancer therapy	Censored

Table 4.1.1-1: Censoring Scheme Used in Primary Definition of PFS

Situation	Date of Progression or Censoring	Outcome
Documented progression per RECIST v1.1 and no new anti-cancer started before	Date of the first documented progression per RECIST v1.1 (excludes clinical progression)	Progressed
No progression and no death, and no new anti-cancer therapy started	Date of last evaluable tumor assessment	Censored
Death without progression per RECIST v1.1 and no new anti-cancer started before	Date of death	Progressed

* Tumor assessments and death if any, occurring after start of subsequent anti-cancer therapy are not considered.

4.1.2 Secondary Definition of Progression Free Survival (Irrespective of Subsequent Therapy)

The secondary definition of PFS (ITT definition) is defined as the time between the date of randomization and the date of first documented tumor progression, based on BICR assessments (per RECIST v1.1 criteria), or death due to any cause, whichever occurs first.

Subjects who die without a reported progression will be considered to have progressed on the date of their death. The following censoring rules will be applied for the secondary definition of PFS:

- Subjects who did not progress or die will be censored on the date of their last evaluable tumor assessment.
- Subjects who did not have any on study tumor assessments and did not die will be censored on their date of randomization.

Censoring rules for the secondary definition of PFS (ITT definition) are presented as follows and in Table 4.1.1-1.

Figure 4.1.2-1: PFS Secondary Definition

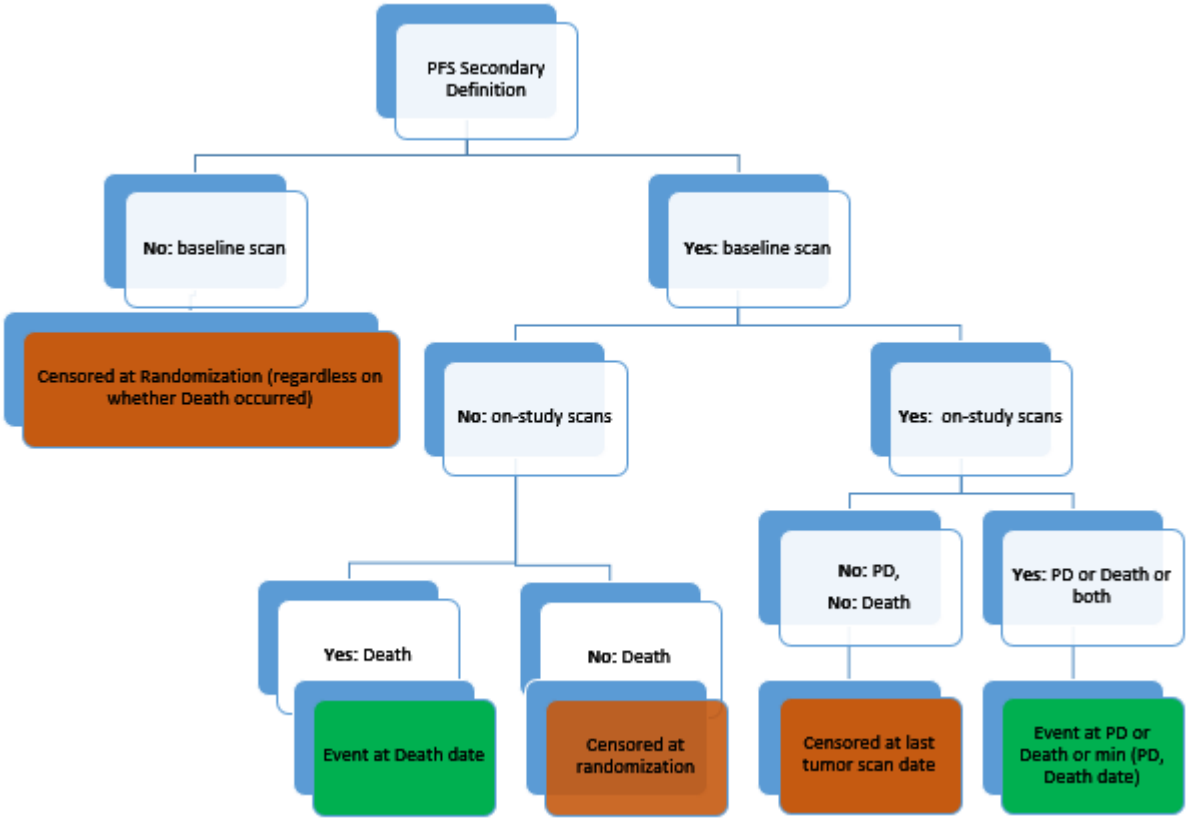


Table 4.1.2-1: Censoring Scheme for Secondary Definition of PFS

Situation	Date of Progression of Censoring	Outcome
No baseline tumor assessment	Date of randomization	Censored
No on-study tumor assessments and no death	Date of randomization	Censored
Documented progression per RECIST v1.1	Date of first documented progression per RECIST v1.1 criteria (excludes clinical progression)	Progressed
No progression and no death	Date of last evaluable tumor assessment	Censored
Death without progression per RECIST v1.1	Date of death	Progressed

Note that the secondary definition will only be used as supportive analysis.

4.2 Secondary Endpoints

4.2.1 Overall Survival

Overall survival (OS) is defined as the time from randomization to the date of death from any cause. For subjects that are alive, their survival time will be censored at the date of last contact date (or "last known alive date"). Overall survival will be censored at the date of randomization for subjects who were randomized but had no follow-up.

Follow-up visit #1 (FU1) should occur 30 days from the last dose and follow-up visit #2 (FU2) occurs approximately 100 days from last dose of study drug. After FU2, survival follow-up will be conducted every 3 months.

4.2.2 Objective Response Rate

Objective Response Rate (ORR) is defined as the number of randomized subjects who achieve a best response of confirmed complete response (CR) or confirmed partial response (PR) based on BICR assessments (using RECIST v1.1 criteria) divided by the number of all randomized subjects. Best Overall Response (BOR) is defined as the best response, as determined by the BICR, recorded between the date of randomization and the date of objectively documented progression per RECIST v1.1 criteria or the date of subsequent therapy (including tumor-directed radiotherapy and tumor-directed surgery), whichever occurs first. For subjects without documented progression or subsequent therapy, all available response designations will contribute to the BOR determination. Confirmation of response is required at least 4 weeks after the initial response.

4.2.2.1 Time to Response

Time to Response (TTR) is defined as the time from randomization to the date of the first confirmed documented response (CR or PR), as assessed by the BICR. TTR will be evaluated for responders (confirmed CR or PR) only.

4.2.2.2 Duration of Response

Duration of Response (DOR) is defined as the time between the date of first confirmed documented response (CR or PR) to the date of the first documented tumor progression as determined by the BICR (per RECIST v1.1 criteria), or death due to any cause, whichever occurs first. Subjects who start subsequent therapy without a prior reported progression will be censored at the last evaluable tumor assessments prior to initiation of the subsequent anti-cancer therapy. Subjects who die without a reported prior progression will be considered to have progressed on the date of their death. Subjects who neither progress nor die, DOR will be censored on the date of their last evaluable tumor assessment. DOR will be evaluated for responders (confirmed CR or PR) only.

4.3 Safety Endpoints

The assessment of safety will be based on the incidence of adverse events (AEs), serious adverse events (SAEs), adverse events leading to discontinuation, adverse events leading to dose modification, select adverse events (select AEs) for EU/ROW Submissions, immune-mediated AEs (IMAEs) for US Submission, other events of special interest (OEOSI), and deaths. The use

of immune modulating concomitant medication will be also summarized. In addition clinical laboratory tests, and immunogenicity (i.e. development of anti-drug antibody) will be analyzed.

4.4 Exploratory Endpoints

4.4.1 Biomarkers

Biomarkers potentially associated with clinical endpoints will be measured by analyzing tumor and blood samples. Biomarker endpoints include, but are not limited to, single-nucleotide polymorphisms (SNPs), proteins in tumor specimens and serum, and immune cell populations.

4.4.1.1 PD-L1 Expression

PD-L1 expression is defined as the percent of tumor cells membrane staining in a minimum of 100 evaluable tumor cells per validated Dako PD-L1 immunohistochemistry (IHC) assay. This is referred to as quantifiable PD-L1 expression. If the PD-L1 staining could not be quantified, it is further classified as:

1) Indeterminate: Tumor cell membrane staining hampered for reasons attributed to the biology of the tumor tissue sample and not because of improper sample preparation or handling.
2) Not evaluable: Tumor tissue sample was not optimally collected or prepared and PD-L1 expression is neither quantifiable nor indeterminate. Not evaluable can be determined from H&E process before the tumor biopsy specimen is sent for PD-L1 evaluation or from the H&E process during PD-L1 evaluation.

Subjects with missing PD-L1 expression are subjects with no tumor tissue sample available for evaluation.

PD-L1 expression will be collected in the IRT as well as in the clinical database. Statistical analysis using PD-L1 expression will be solely based on PD-L1 expression data from clinical database.

Efficacy endpoints defined above (PFS and ORR by BICR, OS) will be analyzed by PD-L1 expression to explore the association of PD-L1 expression on tumor and/or tumor associated immune cells (TAIC) with clinical benefit.

4.4.2 Clinical Outcomes Assessments

The FSKI-19 and EQ-5D-3L patient-reported outcomes will be collected and analyzed in this study.

4.4.2.1 FKSI-19

The NCCN FKSI-19 is a 19-item scale that measures tumor specific HrQoL in kidney cancer subjects. The FKSI-19 uses 5 Likert-type response categories that range from "not at all" to "very much." Subjects are asked to circle the response category that best characterizes their response over the last 7 days on 19 items that include symptoms such as lack of energy, fatigue, appetite, coughing, shortness of breath, pain, nausea, and ability to work.

The instrument yields a total score and four subscale scores: Disease Related Symptoms (DRS), Treatment Side Effects (TSE), and Functional Well Being (FWB).

FKSI-Disease Related Symptoms Physical (FKSI-DRS-P; 12 items; score range 0-48);

FKSI-Disease Related Symptoms Emotional (FKSI-DRS-E; 1 item; score range 0-4);

FKSI-Treatment Side Effects (FKSI-TSE; 3 items; score range 0-12);

FKSI-Functional Well Being (FKSI-FWB; 3 items; score range 0-12);

FKSI-19 Total Score (FKSI-19; 19 items; score range 0-76).

In addition, a score for the FKSI-DRS, a 9-item subscale, can also be calculated.

A higher score indicates fewer symptoms. The recall period is 7 days. Each item is rated on a five-point Likert scale ranging from 0 = "not at all" to 4 = "very much."

If there are missing items, subscale scores can be prorated according to the standard FACT (Functional Assessment of Cancer Therapy) scoring methodology. This can be done using the following formula:

$$\textbf{\textit{Prorated Subscale Score}} = \frac{[\textbf{\textit{Sum of Item Scores}}] \; \textbf{\textit{x}} \; [\textbf{\textit{N of Items in Subscale}}]}{[\textbf{\textit{N of Items Answered}}]}$$

When there are missing data, prorating by subscale in this way is acceptable as long as more than 50% of the items were answered. The total score is then calculated as the sum of the un-weighted subscale scores.

Table 4.4.2.1-1: Time Windows for FKSI-19 and EQ-5D-3L Assessments

Nominal Time-Point (Cycle)	Time Window
Nivolumab+Cabozantinib (Arm A) treatment group	
Baseline (Cycle 1)	Prior to first dose on Day 1
Week 3 (Cycle 2)	Nominal Day 15 (Day 2 thru Day 22,)
Every 2 weeks thereafter (Cycles 3+)	Nominal Days 29+ (+7 days/-6 days, inclusive)
Nivolumab+Ipilimumab+Cabozantinib (Arm B) treatment group	
Baseline (Cycle 1)	Prior to first dose on Day 1
Week 4 (Cycle 2)	Nominal Day 22 (Day 2 thru Day 32, inclusive)
Every 3 weeks up to Week 10 (Cycles 3-4)	Nominal Days 43 and 64 (+10 days/-10 days, inclusive)
Week 13 (Cycle 5)	Nominal Day 85 (+7 days/-10 days, inclusive)
Every 2 weeks thereafter (Cycles 6+)	Nominal Days 99+ (+7 days/-6 days, inclusive)
Sunitinib (Arm C) treatment group	
Baseline (Cycle 1)	Prior to first dose on Day 1
Week 7 (Cycle 2)	Nominal Day 43 (Day 2 thru Day 64, inclusive)
Every 6 weeks thereafter (Cycles 3+)	Nominal Days 85+ (+21 days/-20 days, inclusive)

Table 4.4.2.1-1: Time Windows for FKSI-19 and EQ-5D-3L Assessments

Nominal Time-Point (Cycle)	Time Window
For all treatment groups	
Follow Up:	
Follow-Up 1	30 days from the last dose (± 7 days) or coincide with the date of discontinuation (± 7 days) if date of discontinuation is greater than 42 days after last dose.
Follow-Up 2	100 days (± 7 days) from last dose of study treatment. Subjects must be followed for at least 100 days after last dose of study treatment.
Survival Follow-Up i [a] (i = 1, 2, 3, 4, ...)	3 months (± 14 Days) after Follow-up Visit 2, and subsequent survival follow-up visits every 3 months (± 14 days).

[a] Survival Follow-Up is only relevant for EQ-5D, not FKSI-19 assesments.

4.4.2.2 EuroQoL EQ-5D-3L

Subjects' reports of general health status will be assessed using the EuroQoL Group's EQ-5D-3L. EQ-5D-3L essentially has 2 components: the descriptive system and the visual analogue scale (VAS).

The instrument's descriptive system consists of 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has 3 levels, reflecting "no health problems," "moderate health problems," and "extreme health problems." A dimension for which there are no problems is said to be at level 1, while a dimension for which there are extreme problems is said to be at level 3. Thus, the vectors 11111 and 33333 represent the best health state and the worst health state, respectively, described by the EQ-5D-3L. Altogether, the instrument describes $3^5 = 243$ health states. Empirically derived weights can be applied to an individual's responses to the EQ-5D-3L descriptive system to generate an index measuring the value to society of his or her current health. Such preference-weighting systems have been developed for the UK, US, Spain, Germany, and numerous other populations. For this study, EQ-5D-3L utility index values will be computed using a scoring algorithm based on the United Kingdom Time-Trade-Off (UK TTO) value set[2]

In addition, the EQ-5D-3L includes a VAS, which allows respondents to rate their own current health on a 101-point scale ranging from 0="worst imaginable" health to 100="best imaginable" health state [3].

A change from baseline of 0.08 for the EQ-5D-3L utility index score and of 7 for the EQ-5D-3L VAS are considered minimally important differences for the EQ-5D-3L4.

All questionnaires completed at baseline and on-study will be assigned to a time-point according to the windowing criteria in Table 4.4.2.1-1 and included in the analysis. In case a subject has two

on-study assessments within the same window, the assessment closest to the time-point will be used. And, in the case of two assessments at a similar distance to the time-point, the latest one will be chosen. In the event where the subject has no assessment at all in a specific window, the observation will be treated as missing for that time-point.

4.4.3 Pharmacokinetics

PK will be measured by the serum concentration of nivolumab and/or ipilimumab and/or cabozantinib. Samples will be collected to characterize pharmacokinetics of nivolumab and to explore exposure-safety and exposure-efficacy relationships.

4.4.4 Immunogenicity

Serum samples collected will be analyzed by a validated immunogenicity assay. Selected serum samples may be analyzed by an exploratory orthogonal method that measures anti-nivolumab.

In addition, ad hoc serum samples designated for pharmacokinetic or biomarker assessments may also be used for immunogenicity analysis if required (e.g., insufficient volume for complete immunogenicity assessment or to follow up on suspected immunogenicity related AE).

Further details on immunogenicity background and rationale, definitions, population for analyses and endpoints are described in APPENDIX 3.

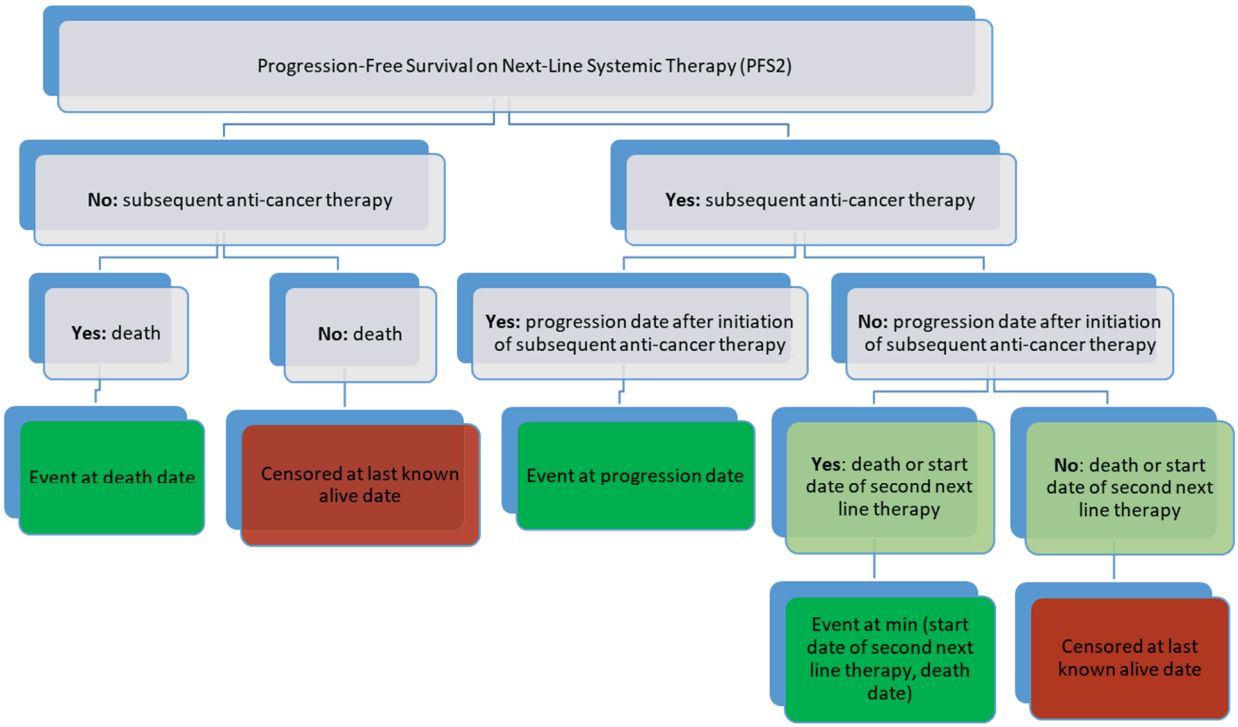
4.4.5 PFS2

PFS on next-line therapy (PFS2) is defined as the time from randomization to objectively documented progression, per investigator assessment, after the next line of therapy or to death from any cause, whichever occurs first. Subjects who were alive and without progression after the next line of therapy will be censored at last known alive date.

The following censoring rules will be applied for PFS2:

- Subjects who did not receive subsequent anti-cancer therapy (i.e. second-line therapy):
  - Subjects who died, the death date is the event date;
  - Else the subject's PFS2 is censored at the last known alive date.
- Subjects who received subsequent anti-cancer therapy (i.e. second-line therapy):
  - Subjects who had a disease progression after the start of subsequent anti-cancer therapy, this disease progression date is the event date;
  - Else if a subject died or start of second next line therapy, the date of min (death, start date of second next line therapy) is the event date;
  - Else the subject's PFS2 is censored at the last known alive date.

Figure 4.4.5-1: PFS2 Definition



5 SAMPLE SIZE AND POWER

The sample size calculations of this study summarized below are based on the target randomized subjects in Arm A and Arm C only. That is, the total randomized subjects will be higher than the target randomized subjects due to randomized subjects into Arm B prior to implementation of CA2099ER Global Revised Protocol 01, which stopped further randomization into Arm B.

The sample size of this study accounts for the primary endpoint of progression-free survival (PFS) per BICR in all randomized subjects. Assuming a 25% screen failure rate, it is expected that approximately 850 subjects will need to be enrolled in order to randomize 638 subjects (319 per arm) in a 1:1 ratio. To represent the normal frequency of the favorable risk group in mRCC, the favorable risk subjects are capped at approximate 25%; thus, at most 212 favorable risk subjects (106 per arm) will be enrolled to randomize 160 favorable risk subjects in a 1:1 ratio. The rest of the enrolled participants will provide approximately 478 intermediate/poor risk randomized subjects (239 per each arm).

The overall alpha for this study is 0.05 (two-sided). PFS will be evaluated for treatment effect at an alpha of 0.05 (two-sided), with at least 95% power. No interim analysis of PFS is planned. OS will be evaluated for treatment effect at an alpha level of 0.05 (two-sided) with 80% power, accounting for two formal interim analyses to assess efficacy.

Sample Size Justification for Primary PFS Endpoint

The primary endpoint of PFS per BICR of Arm A versus Arm C analysis conducted on all randomized subjects. The PFS analysis will occur after approximately 9-10 months minimum

follow-up on all randomized subjects, which will be triggered by approximately 350 events from Arm A and Arm C. The 350 PFS events provide at least 95% power to detect a HR of 0.68 for PFS of Arm A versus Arm C with a type I error of 0.05 (two-sided). The HR of 0.68 corresponds to a 47% increase in the median PFS, assuming a median PFS of 18.2 months for Arm A and 12.4 months for Arm C. It is projected that an observed HR of 0.811 or less, which corresponds to a 2.89 month or greater improvement in median PFS (12.4 versus 15.3 months), would result in a statistically significant improvement in PFS for the Arm A versus Arm C comparison.

If the formal analysis of PFS among all randomized subjects is statistically significant, the formal interim analysis of OS among all randomized subjects will be tested, as per hierarchical testing procedure. If the formal analysis of OS (interim or final, whichever occurs first) among all randomized subjects is statistically significant, then formal analysis of ORR among all randomized subjects will be tested, as per hierarchical testing procedure. The formal ORR analyses specified in the SAP supersede those specified in the protocol.

Sample Size Computation for Secondary OS Endpoint

The secondary endpoint of OS in all randomized subjects specifies the comparison of Arm A versus Arm C. Among all randomized subjects, approximately 254 events (ie, deaths) in Arm A and Arm C provides at least 80% power to detect a HR of 0.70 for OS of Arm A and Arm C with an overall type 1 error of 0.05 (two-sided) for each test. The HR of 0.70 corresponds to a 43% increase in the median OS, assuming a median OS of 47.1 months for Arm A and 33 months for Arm C.

Two formal interim analyses of OS are planned for this study. The first interim analysis is planned at the time of final PFS analysis and it is expected to observe 165 OS events (65% of the targeted OS events for final analysis) and the second interim analysis is planned to occur after observing approximately 211 events (83% of targeted OS events needed for final analysis). The stopping boundaries at interim and final analyses will be derived based on the number of deaths using O'Brien and Fleming α spending function. For example, with 165, 211, and 254 observed events in Arm A and Arm C at the first interim, second interim, and final analyses, the respective stopping boundaries would be α=0.011 (two-sided), α=0.025 (two-sided), and α=0.041 (two-sided). If the first interim analysis is performed exactly at 165 deaths, it is projected that an observed HR of 0.673 or less, which corresponds to a 16.0 month or greater improvement in median OS (33 versus 49 months), would result in a statistically significant improvement in OS for the Arm A versus Arm C comparison. At the second interim analysis with 211 deaths, it is projected that an observed HR of 0.734 or less, which corresponds to a 12.0 month or greater improvement in median OS (33 versus 45 months), would result in a statistically significant improvement in OS for the Arm A versus Arm C comparison. At the time of final OS analysis when there are 254 deaths, it is projected that an observed HR of 0.774 or less, which corresponds to a 9.6 month or greater improvement in median OS (33 versus 42.6 months), would result in a statistically significant improvement in OS for the Arm A versus Arm C comparison.

Assuming a constant accrual rate (an average rate of 3 subjects/month in the first 4 months, afterwards an average rate of 42 subjects/month), the accrual will take approximately 19 months. The final PFS analysis will not occur prior to these conditions being met:

- at least 8 months minimum follow-up on all randomized subjects;
- at least 283 PFS events, which provide at least 90% power to detect a HR of 0.68 for PFS of Arm A versus Arm C; and
- at least 149 OS events, which provide 66% power if the observed HR for OS was 0.60.
  (Note that if the analysis of first interim analysis OS takes place with 149 OS events, the alpha spending for the OS comparison would be 0.007 with a critical HR=0.643.)

This expected PFS analysis will occur at approximately 29 months from FPFV. The second interim and final analyses of OS are expected to occur approximately 34 months and 40 months from FPFV, respectively. Table 5-1 summarizes the results of these calculations.

Table 5-1: Summary of Sample Size Parameters and Schedule of Analyses

Primary/Secondary Endpoints	PFS (Primary)	OS (Secondary)
Primary analysis population	All Randomized Subjects	
Accrual rate per month for all randomized population	3 subjects/month in the first 4 months, afterwards an average rate of 42 subjects/month	
Power	95%	80%
Alpha	0.05 2-sided	0.05 2-sided (0.011 at IA1, 0.025 at IA2, 0.041 at FA)
Hypothesized median control vs exp (months)	12.4 vs 18.2	33 vs 47.1
Hypothesized hazard ratio	0.68	0.70
Critical hazard ratio (observed hazard ratio at which a statistically significant difference would be observed) / Difference in median (months) Corresponding to a minimal clinically significant effect size (FA)	0.811 / 2.89	0.774 / 9.6
Critical HR at interim analysis-1(IA1) /effect size	N/A	0.673 / 16.0
Expected number of event for IA1 (percentage of target events)	N/A	165 (65%)
Timing of IA1 from FPFV (months)	N/A	29
Critical HR at interim analysis-2(IA2) /effect size	N/A	0.734 / 12.0
Expected number of event for IA2 (percentage of target events)	N/A	211 (83%)
Timing of IA2 from FPFV (months)	N/A	34
Accrual duration (months)	19	19

Table 5-1: Summary of Sample Size Parameters and Schedule of Analyses

Primary/Secondary Endpoints	PFS (Primary)	OS (Secondary)
Timing of final analysis (FA) from FPFV (months)	29	40
Sample size	638	638
Target number of events (Event Goal)	350	254

6 STUDY PERIODS, TREATMENT REGIMENS AND POPULATIONS FOR ANALYSES

6.1 Study Periods

- Baseline period:
  - Baseline evaluations or events will be defined as evaluations or events that occur before the date and time of the first dose of study treatment. Evaluations (laboratory tests, pulse oximetry and vital signs) on the same date and time of the first dose of study treatment will be considered as baseline evaluations. Events (AEs) on the same date and time of the first dose of study treatment will not be considered as pre-treatment events.
  - In cases where the time (onset time of event or evaluation time and dosing time) is missing or not collected, the following definitions will apply:
    - Pre-treatment AEs will be defined as AEs with an onset date prior to but not including the day of the first dose of study treatment;
    - Baseline evaluations (laboratory tests, pulse oximetry and vital signs) will be defined as evaluations with a date on or prior to the day of first dose of study treatment.
  - If there are multiple valid assessments on or prior to the first dose of study treatment:
    - For laboratory tests, the latest non missing labs value on or before first dose date (and time if collected) will be used as the baseline in the analyses. For 'LIPASE' and 'GLUCOSE', for treated subjects only, the last predose assessment with non-missing toxicity grade will be considered as baseline. If multiple assessments exist with the same collection date (and time if collected) and entry date and time, then the first observation is used as baseline.
    - For PD-L1, among the records prior to or on first dose date (and time if collected), identify first those with quantifiable test result. If there are no records with quantifiable test result, then select those with indeterminant result ("INDETERMINATE"). If there are no records with indeterminant test result, then select those with unavailable result ("NOT EVALUABLE). If there are no records with unavailable test result, then select those with not reported or not available result (all other records). The latest record will be used as the baseline in the analyses. If there is more than one record for the latest date, then choose the one with the greatest specimen ID.
    - For Anti-Drug Antibody (ADA), the record related to the most recent assessment among those records where date (and time if collected) of Nivolumab immunoglobulin (IMG) assessment is less than or equal to the date (and time if collected) of the first Nivolumab dose date.

- Post baseline period:
  - On-treatment AEs will be defined as AEs with an onset date and time on or after the date and time of the first dose of study treatment (or with an onset date on or after the day of first dose of study treatment if time is not collected or is missing). For subjects who are off study treatment, AEs will be included if event occurred within a safety window of 30 days (or 100 days depending on the analysis) after the last dose of study treatment. No "subtracting rule" will be applied when an AE occurs both pre-treatment and post-treatment with the same preferred term and grade.
  - On-treatment evaluations (laboratory tests, pulse oximetry and vital signs) will be defined as evaluations taken after the day (and time, if collected and not missing) of first dose of study treatment. For subjects who are off study treatment, evaluations should be within a safety window of 30 days (or 100 days depending on the analysis) after the last dose of study treatment.
- Late-emergent drug-related AEs will be defined as drug-related AEs with an onset date greater than 100 days after the last dose of study treatment in subjects who are off study treatment.

6.2 Treatment Regimens

The treatment group "**as randomized"** will be retrieved from the IRT system:

- Arm A: Experimental arm: nivolumab + cabozantinib
- Arm B: Experimental arm: nivolumab + ipilimumab + cabozantinib
- Arm C: Control arm: sunitinib

The treatment group "**as treated**" will be the same as the arm as randomized by IRT. However, if a subject received the incorrect drug for **the entire period** of treatment, the subject's treatment group will be defined as the incorrect drug the subject actually received.

Unless otherwise specified, the safety analysis will be based on the treatment group "as treated".

Unless otherwise specified, the efficacy analysis will be based on the treatment group "as randomized".

6.3 Populations for Analyses

All analyses will be performed using the treatment arm as randomized (intent to treat), with the exception of dosing and safety, for which the treatment arm as received will be used. For purposes of analysis, the following populations are defined in Table 6.3-1, and all populations for analyses given in this table refer to those subjects in Arm A and Arm C. Those subjects who randomized to Arm B prior to Revised Protocol 01 will be considered as part of the population of interest for descriptive summary of efficacy and safety analyses.

Table 6.3-1: Populations for Analyses

Population	Description
All Enrolled Subjects	All subjects who sign informed consent and were registered into the IRT.
All Randomized Subjects	All subjects who were randomized will be used for analyses of demography, protocol deviations, baseline characteristics, primary efficacy analysis, secondary efficacy analyses, and outcome research analysis which will be performed for this population.
All Treated Subjects	All subjects who received at least one dose of any study medication. This is the primary population for exposure and safety analyses.
Intermediate/Poor Risk Subjects	All subjects who were randomized with baseline IMDC prognostic score ≥1 at the time of randomization (per IRT). This population will be used for subset analyses of demography, protocol deviations, baseline characteristics, primary efficacy analysis, and secondary efficacy analyses on intermediate/poor risk subjects.
All Intermediate/Poor Risk Treated Subjects	All intermediate/poor risk subjects who received any dose of study therapy. This population will be used for subset analyses of exposure and safety analyses on intermediate/poor risk subjects.
Pharmacokinetic Subjects	All subjects with available serum time-concentration data from randomized subjects dose with nivolumab and cabozantinib.
Immunogenicity Subjects	All subjects with available data from randomized subjects dose with nivolumab and cabozantinib.
PD-L1 Treated Subjects	All subjects with a PD-L1 assessment at baseline who received any dose of study therapy.

7 STATISTICAL ANALYSES

7.1 General Methods

Unless otherwise noted, discrete variables will be tabulated by the frequency and proportion of subjects falling into each category, grouped by treatment. Percentages given in these tables will be rounded to the first decimal and, therefore, may not always sum to 100%. Percentages less than 0.1 will be indicated as '< 0.1'. Continuous variables will be summarized by treatment group using the mean, standard deviation, median, minimum, and maximum values.

Time-to-event variables (e.g. time-to resolution) will be analyzed using the Kaplan-Meier technique. When specified, the median will be reported along with 95% CI using Brookmeyer and Crowley method[5] (using log-log transformation for constructing the confidence intervals[6]).

Unless otherwise specified, the stratified log-rank test will be performed to test the comparison between time to event distributions. Unless otherwise specified, the stratified hazard ratio between 2 groups along with CI will be obtained by fitting a stratified Cox model with the group variable as a unique covariate.

Confidence intervals for binomial proportions will be derived using the Clopper-Pearson method.[7] The unweighted difference in ORRs between the two treatment arms and corresponding asymptotic 95% CI will be estimated using a Newcombe method.[8]

P-values from sensitivity analyses for efficacy endpoints are for descriptive purpose only and there will be no multiplicity adjustment for these analyses.

The conventions to be used for imputing missing and partial dates for analyses requiring dates are described in Section 8.

Note that, the outputs will present Arm A, Arm B, and Arm C in the following way:

- All analysis specified in Section 7.3 and 7.4 will present A, B, and C together.
- For purposes of efficacy analyses, all randomized subjects in Arm A and Arm C will be tabulated and presented in the same table. A few seperate efficacy summary tables will be generated for those subjects randomized to Arm B, which will be specified in the corresponding sections.
- For purposes of safety analyses, all treated subjects in Arm A and Arm C will be tabulated and presented in the same table. A few separate safety summary tables will be generated for those subjects treated in Arm B, which will be specified in the corresponding sections.
- Summaries of outcome research and biomarker data will be tabulated on the Arm A and C subjects only.

7.1.1 Adverse Events, Serious Adverse Events, Multiple events, Select Adverse Events, Other Events of Special Interest and Immune-Mediated Adverse Events

Drug-related AEs are those events with relationship to study drug "Related", as recorded on the CRF. If the relationship to study drug is missing, the AE will be considered as drug-related.

Serious adverse events consist of AEs deemed serious by the Investigator and flagged accordingly in the CRF and clinical database.

Adverse events leading to study drug discontinuation are AEs with action taken regarding study drug(s) = "Drug was discontinued".

Adverse events leading to dose delay are AEs with action taken regarding study drug(s) = "Drug was delayed".

Adverse event that led to dose delay of the oral drug (similarly defined as dose omission or dose interruption) will be coded with action "Drug was interrupted".

Adverse events leading to dose reduction are AEs with action taken regarding study drug(s) = "Dose was reduced".

Adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA), and the most recent version of the dictionary at the time of the database lock will be used. Adverse events results will be graded for severity using NCI Common Terminology Criteria for Adverse

Events (CTCAE) and the most recent version of the criteria at the time of the database lock will be used.

In the AE summary tables, unless otherwise specified, subjects will be counted only once at the Preferred Term (PT), only once at the System Organ Class (SOC), and only once at subject level for the counting of total number of subjects with an AE. The AE tables will be sorted by the SOCs and then PTs. SOC will be ordered by descending frequency overall and then alphabetically. PTs will be ordered within SOC by descending frequency overall and then alphabetically. The sorting will be done based on the 'Any Grade' column of the experimental arm when arms are presented side-by-side.

Unless otherwise specified, the AE summary tables will be restricted to on-treatment events regardless of the causality.

Analyses that take into account the multiple occurrences of a given adverse event will be conducted (see Section 7.6.9). To prepare these analyses, the CRF data will be processed according to standard BMS algorithms[9] in order to collapse adverse event records into unique records based on the preferred term. These data will be presented as the rate per 100 person-years of exposure. These analyses will take into account all on-treatment events (allowing more than 1 event per subject) and the total exposure time. The person-year exposure will be computed as the sum over the subjects' exposure expressed in years where the exposure time is defined as

- (Date of last dose of study treatment - date of first dose of study treatment + 31 days (or 101 days, depending on the analysis))/365.25, for subject who are off study treatment and were followed for at least 30 days (or 100 days, depending on the analysis) after last dose of study treatment.
- (Last known alive date - date of first dose of study treatment +1)/365.25, for subjects who are still on-treatment or who are off study treatment and were followed less than 30 days (or 100 days depending on the analysis) after last dose of study treatment.

7.1.1.1 Select Adverse Events (EU/ROW Submissions)

The select Adverse Events (select AEs) consist of a list of preferred terms grouped by specific category (e.g. pulmonary events, gastrointestinal events categories, etc.). AEs that may differ from or be more severe than AEs caused by non-immunotherapies and AEs whose early recognition and management may mitigate severe toxicity are included as select AEs. Categories of select AEs may include subcategories (e.g. adrenal disorders, diabetes, pituitary disorders, and thyroid disorders are subcategories of the endocrine event category).

The list of MedDRA preferred terms used to identify select adverse events is revisited quarterly and updated accordingly. The preferred terms used for the selection at the time of the database lock will be provided by categories/subcategories.

In addition to the frequency and worst severity of select AEs, time-to onset, time-to resolution, and time-to resolution where immune modulating medication was initiated will be analyzed for each specific category/subcategory of drug-related select AEs when applicable.

Further details on the definitions time-to onset and time-to resolution are described in APPENDIX 1.

7.1.1.2 Other Events of Special Interest

Other events of special interest (OEOSI) consist of a list of preferred terms grouped by specific category (e.g. Myositis Event, Myocarditis Event, Demyelination Event, Guillain-Barre Syndrome, Pancreatitis Event, Uveitis Event, Encephalitis Event, Myasthenic Syndrome, Rhabdomyolysis Event, Graft Versus Host Disease). The list of MedDRA preferred terms used to identify OEOSI is revisited quarterly and updated accordingly. The preferred terms used for the selection at the time of the database lock by categories will be provided.

7.1.1.3 Immune-Mediated Adverse Events (US Submission)

In order to further characterize AEs of special clinical interest, analysis of immune-mediated AEs (IMAE) will be conducted. IMAEs are specific events (or groups of PTs describing specific events) that include pneumonitis, diarrhea/colitis, hepatitis, nephritis/renal dysfunction, rash, endocrine (adrenal insufficiency, hypothyroidism/thyroiditis, hypothyroidism, thyroiditis, hyperthyroidism, diabetes mellitus, and hypophysitis), and other specific events, considered as potential immune-mediated events by investigator that meet the definition summarized below:

- those occurring within 100 days of the last dose,
- regardless of causality,
- treated with immune-modulating medication (of note, endocrine AEs such as adrenal insufficiency, hypothyroidism/thyroiditis, hypothyroidism, thyroiditis, hyperthyroidism, diabetes mellitus, and hypophysitis are considered IMAEs regardless of immune-modulating medication use, since endocrine drug reactions are often managed without immune-modulating medication).
- with no clear alternate etiology based on investigator assessment, or with an immune-mediated component

The list of MedDRA preferred terms used to identify IMAEs is revisited quarterly and updated accordingly. The preferred terms used for the selection at the time of the database lock by categories will be provided.

7.1.2 Laboratory Tests

Clinical laboratory parameters (hematology, serum chemistry and electrolytes) will be evaluated.

Laboratory tests will be graded using the NCI Common Terminology Criteria, and the most recent version of the criteria at the time of the database lock will be used.

Clinical laboratory data will be first analyzed using International System of Units (SI).

Analyses will be repeated using US conventional units.

In the laboratory summary tables, unless otherwise specified, subjects will be counted only once for each lab parameter according to their worst on treatment CTC grade (worst being the highest

CTC grade). The laboratory tables and listings will be sorted by laboratory category, laboratory subcategory and laboratory test code sequence number.

7.1.3 Immunogenicity Data

Blood samples for immunogenicity analysis will be collected from subjects assigned to the experimental treatment group(s) according to the protocol schedule. Samples will be evaluated for development of Anti-Drug Antibody (ADA) by a validated electrochemiluminescent (ECL) immunoassay.

7.2 Study Conduct

The following programmable deviations will be considered as relevant protocol deviations and summarized by treatment group and overall in all randomized subjects. Non-programmable relevant eligibility and on-treatment protocol deviations, as well as significant (both programmable and non-programmable) eligibility and on-treatment protocol deviations will be reported through ClinSIGHT listings.

Eligibility:

- Subjects with baseline KPS < 70%
- Subjects who received prior systemic anti-cancer treatment in the metastatic setting
- Subjects without histologically confirmed RCC with a clear-cell component, documented advanced or metastatic RCC

On-study:

- Subjects receiving anti-cancer therapy (chemotherapy, hormonal therapy, immunotherapy, standard or investigational agents for treatment of cancer) while on study therapy
- Subjects treated differently than as randomized (subjects who received the wrong treatment, excluding the never treated)

Enrollment by country and site, and enrollment by month will be summarized and listed for all enrolled subjects.

A by-subject listing of batch numbers for all treated subjects will be provided.

7.3 Study Population

Analyses in this section will be tabulated for all randomized subjects by treatment group as randomized, unless otherwise specified.

7.3.1 Subject Disposition

The total number of subjects enrolled (randomized or not randomized) will be presented along with the reason for not being randomized. This analysis will be performed on the all enrolled subjects population only.

Number of subjects randomized but not treated along with the reason for not being treated will be tabulated by treatment group as randomized.

Number of subjects who discontinued study treatment along with corresponding reason will be tabulated by treatment group as treated. Reason for discontinuation will be derived from subject status CRF page. This analysis will be performed only on the all treated subjects population.

A by-subject listing for all treated subjects will be provided showing the subject's off treatment date and whether the subject continue in the study along with the reason for going off study. A by-subject listing for all enrolled subjects will also be provided, showing whether the subject was randomized along with the reason for not being randomized.

Note that for all intermediate/poor risk treated subjects in Arm A and Arm C, number of subjects who discontinued study treatment along with corresponding reason will be tabulated by treatment group as treated.

7.3.2 Demographics and Other Baseline Disease Characteristics

The following demographic and baseline disease characteristics will be summarized and listed by treatment group as randomized:

- Age
- Age categorization (< 65, ≥ 65 and < 75, ≥ 75 and < 85, ≥ 85, ≥ 75, ≥ 65)
- Sex (Male, Female)
- Race
- Region (Region (US/Canada/W.Europe/N.Europe vs. ROW) (source: IRT)
- Ethnicity (Hispanic/Latino and Not Hispanic/Latino)
- Karnofsky performance status (70, 80, 90, 100)
- Baseline IMDC prognostic score (0, 1-2, ≥ 3) (source: IRT)
- Baseline IMDC prognostic score (0, 1-2, ≥ 3) (source: CRF)
- Time from initial disease diagnosis to randomization (<1 year, ≥1 year)
- Baseline LDH level (≤ 1.5 x ULN, >1.5 x ULN)
- Hemoglobin (<LLN, ≥ LLN)
- Corrected Calcium (≤ 10 mg/dl, >10mg/dl)
- Absolute Neutrophil Count (≤ ULN, > ULN)
- Platelet Count (≤ ULN, > ULN)
- Baseline Alkaline phosphatase (< ULN, ≥ ULN)
- Prior nephrectomy (Yes, No)
- Prior radiotherapy (Yes, No)
- Baseline PD-L1+ status based on a 1% cut off (≥ 1% vs. < 1% or indeterminate)
- Baseline PD-L1+ status based on a 5% cut off (≥ 5% vs. < 5% or indeterminate)
- Baseline PD-L1+ status based on a 10% cut off (≥ 10% vs. < 10% or indeterminate)

- Number of disease sites per subject (1, 2, 3, 4 , >4)
- Tumor burden: sum of the diameters of target lesions at baseline
- Most common sites of metastasis
- Sarcomatoid features (Yes, No)
- Stage at the initial diagnosis (Stage IV, non-Stage IV)
- Pre-treatment events tumor assessment (per Investigator)

Summary table (cross-tabulation) by treatment group for stratification factor (except for region) will be provided to show any discrepancies between what was reported through IRT vs. CRF at baseline. This summary will be performed based on all randomized subjects.

- IMDC Prognostic Score: 0 versus 1-2 versus 3-6
- Region: US/Canada/W Europe/N Europe versus ROW
- PD-L1 tumor expression: ≥ 1% versus < 1% or indeterminate

A listing of randomization scheme presenting randomized treatment group and as treated treatment group will be provided for all randomized subjects.

Note that for all intermediate/poor risk subjects in Arm A and Arm C, demographic and baseline disease characteristics will be summarized and listed by treatment group as randomized.

7.3.3 Medical History

A by-subject listing of general medical history for all randomized subjects will be provided.

7.3.4 Prior Therapy Agents

Prior adjuvant or neo-adjuvant therapy will be summarized by treatment group and overall.

Prior systemic cancer therapy will be summarized by treatment group and overall and listed by subject.

Prior radiotherapy and prior surgery related to cancer will be listed by subject.

7.3.5 Physical Examinations

Subjects with abnormal baseline physical examination will be listed by subject.

7.3.6 Baseline Physical Measurements

Baseline physical measurements will be listed by subject.

7.4 Extent of Exposure

Listings will include all available exposure data. Analyses will be performed by treatment group "as treated" in all treated subjects, unless otherwise specified.

7.4.1 Administration of Study Therapy

The following parameters will be summarized (descriptive statistics) by study therapy and treatment group:

- Number of doses received
- Cumulative dose
- Relative dose intensity (%) using the following categories: < 50%; 50 - < 70%; 70 - < 90%; 90 - < 110%; ≥ 110%
- Average daily dose

Duration of study therapy will be summarized (descriptive statistics) by treatment group.

A by-subject listing of dosing of study medication (record of study medication, infusion details, and dose changes) will be also provided.

Note that similar study therapy table will be summarized for all intermediate/poor risk treated subjects in Arm A and Arm C.

Table 7.4.1-1: Study Therapy Parameter Definitions for Arm A and C

	Nivolumab	**Cabozantinib**	**Sunitinib**
Dosing Schedule per Protocol	240 mg every 2 weeks	40 mg PO once daily	50 mg PO once daily for 4 weeks followed by 2 weeks off.
Dose	mg	mg	mg
Cumulative Dose	mg sum of the doses administered to a subject	mg sum of the doses administered to a subject	mg sum of the doses administered to a subject
Relative Dose Intensity (%)	[Cum dose (mg) /((Last dose date– First dose date + 14) × 240/14)] × 100	See below	See below
Duration of Treatment	*Last dose date - Start dose date +1*	*Last dose date - Start dose date + 1*	*Last dose date - Start dose date +15*

Additional Parameters -Cabozantinib treatment

Average daily dose (in mg/day) is defined as:

Sum of all Cabozantinib doses in mg actually received / duration of treatment in days.

Since Cabozantinib treatment consists of 40 mg PO daily dose, the planned dose intensity of Cabozantinib is 40 mg/day.

Relative dose intensity for Cabozantinib (%) is defined as: *(Average daily dose / 40) x 100.*

Additional Parameters -Sunitinib treatment

Average daily dose (in mg/day) is defined as:

Sum of all Sunitinib doses in mg actually received / duration of treatment in days

Since Sunitinib treatment consists of 50 mg PO daily dose for 4 weeks followed by 2 weeks of washout period, the planned dose intensity of Sunitinib is 33.33 mg/day (50 mg x 28 days / 42 days).

Relative dose intensity for Sunitinib (%) is defined as: *(Average daily dose / 33.33) x 100.*

Table 7.4.1-2: Study Therapy Parameter Definitions for Arm B (Cycle 1-4)

	Nivolumab	Ipilimumab	Cabozantinib
Dosing Schedule per Protocol	3 mg/kg every 3 weeks for 4 doses	1 mg/kg every 3 weeks for 4 doses	40 mg PO once daily
Dose	Dose (mg/kg) is defined as Total Dose administered (mg)/Most recent weight (kg). Dose administered in mg at each dosing date and weight are collected on the CRF.	Dose (mg/kg) is defined as Total Dose administered (mg)/Most recent weight (kg). Dose administered in mg at each dosing date and weight are collected on the CRF.	mg
Cumulative Dose	Cum Dose (mg/kg) is the sum of the doses administered to a subject.	Cum Dose (mg/kg) is the sum of the doses administered to a subject.	mg sum of the doses administered to a subject
Cycle Duration$_{(i)}$ (wk)	(Dose date$_{(i+1)}$ - Dose date$_{(i)}$)/7	(Dose date$_{(i+1)}$ - Dose date$_{(i)}$)/7	N/A
Cycle Intensity$_{(i)}$ (mg/kg/wk)	Dose$_{(i)}$/Cycle Duration$_{(i)}$	Dose$_{(i)}$/Cycle Duration$_{(i)}$	N/A
Relative Cycle Intensity $_{(i)}$ (%)	(Cycle Intensity$_{(i)}$/intended dose per week)(i) * 100	(Cycle Intensity$_{(i)}$/intended dose per week)(i) * 100	N/A
Relative Dose Intensity (%)	Sum of all Relative Cycle Intensities divided by N	Sum of all Relative Cycle Intensities divided by N	(Average daily dose / 40) x 100
Duration of Treatment	*Last dose date - Start dose date +1*	*Last dose date - Start dose date +1*	*Last dose date - Start dose date +1*

Table 7.4.1-3: Study Therapy Parameter Definitions for Arm B (Cycle 5 Onward)

	Nivolumab	Cabozantinib
Dosing Schedule per Protocol	240 mg every 2 weeks	40 mg PO once daily
Dose	mg	mg
Cumulative Dose	mg sum of the doses administered to a subject	mg sum of the doses administered to a subject
Relative Dose Intensity (%)	[Cum dose (mg) /((Last dose date– First dose date + 14) × 240/14)] × 100	(Average daily dose / 40) x 100
Duration of Treatment	*Last dose date - Start dose date +1*	*Last dose date - Start dose date + 1*

7.4.2 Modifications of Study Therapy

7.4.2.1 Dose Delays

Each nivolumab infusion or sunitinib dose may be delayed. A dose will be considered as actually delayed if the delay is exceeding 3 days (ie greater than or equal to 4 days from scheduled dosing date) for nivolumab. Reason for dose delay will be retrieved from CRF dosing pages. It is worth noting that during the two week mandatory washout period for sunitinib, a daily dose of 0 mg will be entered in the CRF pages, with corresponding reason for dose modification recorded as "No Change".

If cabozantinib is given daily, a daily dose of 0 mg entered in the CRF pages will be considered as delay. If cabozantinib is given every other day, then a daily dose of 0 mg will be entered every other day in the CRF pages, with corresponding reason for dose modification recorded as "No Change". If there are more than one 0 mg daily dose entered consecutively, then this will be considered as delay.

The following parameters will be summarized by treatment group:

- Number of subjects with at least one dose delayed, the number of dose delays per subject, the reason for dose delay and the length of dose delay.

Note that similar dose delay summary table for Arm A and Arm C will be summarized for all intermediate/poor risk treated subjects.

For Arm B, both nivolumab and ipilimumab can be delayed at the same cycle. A dose will be considered as actually delayed if the delay is exceeding 3 days (ie greater than or equal to 4 days from scheduled dosing date) for nivolumab and ipimumab. Cabozantinib a daily dose of 0 mg entered in the CRF pages will be considered as delay. Similar table will be produced for Arm B

subjects previously randomized to Arm B continue with Arm B treatment and continue with Arm B clinical planned events, per protocol.

7.4.2.2 Infusion Interruptions and Rate Changes

Each nivolumab or ipilimumab infusion can be interrupted and/or the IV infusion rate can be reduced. This information will be retrieved from CRF dosing pages.

The following parameters will be summarized by treatment group:

- Number of subjects with at least one dose infusion interruption, the reason for interruption, and the number of infusion interruptions per subject.
- Number of subjects with at least one IV infusion rate reduction, the reason for reduction and the number of infusion with IV rate reduction per subject.

Note that similar summary table for Arm A and Arm C will be summarized for all intermediate/poor risk treated subjects.

7.4.2.3 Dose Escalations

Dose escalations are permitted for cabozantinib and sunitinib but not for nivolumab and ipilimumab.

7.4.2.4 Dose Reductions

Dose reductions are permitted for cabozantinib and sunitinib but not for nivolumab and ipilimumab.

Dose reduction for subjects treated with sunitinib is defined as at least one day with a non zero dose smaller than 50 mg and smaller than previous non zero dose with a CRF reason different from "Dosing Error" or "No Change".

Dose reduction for subjects treated with cabozantinib is defined as at least one day with 20 mg or 20 mg every other day with a CRF reason different from "Dosing Error" or "No Change".

Note that similar dose redcution summary table for Arm A and Arm C will be summarized for all intermediate/poor risk treated subjects.

The following summaries will be presented for the cabozantinib component of study treatment:

i) For dose reductions due to AE
   Categorical summaries for:

   - Subjects with any dose reduction
   - Dose levels received by a subject
   - Lowest non-zero dose level received
   - Last non-zero dose level received
   - Last dose level received (including dose holds)

Descriptive statistics for:

- Duration of treatment in months for each dose level (40 mg, 20 mg, 0 mg)
- Time to second dose level reduction (first receipt of 20mg) (days)

ii) Summaries for dose holds due to AE (those with 0 mg due to AE):

- Descriptive statistics for number of dose holds due to an AE
- Descriptive statistics for duration of dose holds per dose hold and per subject due to an AE, calculated as (stop date of hold – start date of hold + 1)
- Categorical summary for subjects with duration of holds due to an AE that can be classified as any number of days, ≥ 7 days, ≥ 14 days, ≥ 21 days, and >42 days
- Descriptive statistics for time to first dose hold, time to first dose hold that ≥ 7 days, ≥ 14 days, ≥ 21 days, and >42 days. The time to dose hold is calculated as (start date of the hold – first dose date + 1)
- Descriptive statistics for time to second dose hold, time to second dose hold that was ≥ 7 days, ≥ 14 days, ≥ 21 days, and >42 days.

iii) Summaries for dose modifications (defined as a reduction or hold) due to AE:

- Frequency counts and percentages for subjects with any dose modifications
- Descriptive statistics for number of dose modifications (0-3)
- Descriptive statistics for time to the first dose modification
- Descriptive statistics for time to the second dose modification

7.4.2.5 Dose Omissions

Dose omissions are not permitted.

7.4.3 Concomitant Medications

Concomitant medications, defined as medications other than study medications which are taken at any time on-treatment (i.e. on or after the first day of study therapy and within 100 days following the last dose of study therapy), will be coded using the UMC WHO Drug Global Dictionary.

The following summary table will be provided:

- Concomitant medications (subjects with any concomitant medication, subjects by medication class and generic term)

A by-subject listing will accompany the table.

7.4.3.1 Immune Modulating Medication

Immune modulating concomitant medications are medications entered on an immune modulating medication form or available from the most current pre-defined list of immune modulating medications. The list of anatomic class, therapeutic class and generic name used for the selection at the time of the database lock will be provided.

The percentage of subjects who received immune modulating concomitant medication for

- management of adverse event
- premedication
- other use
- any use
- management of drug-related select adverse event (any grade, grade 3-5) by select AE category/ subcategory (EU/ROW Submissions)
- management of IMAEs (any grade, grade 3-5) by IMAE category (US Submission) will be reported separately for each treatment group (percentages of treated subjects by medication class and generic term).

For each category/subcategory of drug-related select AEs (any grade, grade 3-5) and IMAEs (any grade, grade 3-5), the following will be reported for each treatment group:

- The total immune modulating medication treatment duration (excluding overlaps), duration of high dose of corticosteroid, initial dose of corticosteroid, and tapering duration (summary statistics)

Duration represents the total duration the subject received the concomitant medication of interest. If the subject took the medication periodically, then DURATION in the summation of all use. Initial dose represents the dose of the concomitant medication of interest received at the start of the event. In the case multiple medications started on the same date, the highest equivalent dose is chosen and converted to mg/kg by dividing by the subject's recent weight.

These analyses, except the ones related to IMAEs will be conducted using the 30-day safety window. The analyses related to IMAEs will be conducted using the 100-day safety window.

7.4.3.2 Subsequent Cancer Therapy

Number and percentage of subjects receiving subsequent cancer therapies will be summarized for all randomized subjects. Categories include:

- Subsequent systemic therapy
- Subsequent surgery for treatment of tumors
- Subsequent radiotherapy for treatment of tumors

A by-subject listing of subsequent cancer therapy will also be produced for all randomized subjects.

Note that similar tables will be summarized for all intermediate/poor risk subjects in Arm A and Arm C.

7.5 Efficacy

Analyses in this section will be tabulated for all randomized subjects in Arm A and Arm C, unless otherwise specified. A few seperate efficacy summary tables will be generated for those subjects who randomized to Arm B prior to Revised Protocol 01, which is specified in the relevant subsections below.

Principal analyses of progression free survival (PFS) and objective response rate (ORR) will be based on the Blinded Independent Central Review (BICR) evaluation, unless noted otherwise.

Analyses in this section will be tabulated for all randomized subjects by treatment group as randomized, unless otherwise specified.

Unless stated otherwise, whenever a stratified analysis is specified, the following stratifications factors (recorded at randomization as per IRT) will be used:

- IMDC Prognostic Score: 0 versus 1-2 versus 3-6
- Region: US/Canada/W Europe/N Europe versus ROW
- PD-L1 tumor expression: ≥ 1% versus < 1% or indeterminate

The key secondary objective Overall Survival among all randomized subjects will be tested after conducting the primary objective analyses of PFS on all randomized subjects. For assessing this secondary objective of this study, a hierarchical testing procedure[10] will be used so that the overall experiment-wise Type I error rate is two-sided 0.05.

The secondary objective of ORR among all randomized subjects will be tested after conducting the key secondary objective analyses of OS on all randomized subjects, per a hierarchical testing procedure so that the overall experiment-wise Type I error rate is two-sided 0.05. Note that the formal ORR analysis specified in the SAP supersede those specified in the protocol.

Confidence intervals (CI) for primary and secondary endpoint analyses included in hierarchy will be based on nominal significance level adjusted for primary endpoints and interim analyses to preserve overall type one error rate.

Alpha (α) for the CI will be the same as nominal significance level for hypothesis testing. CIs for other endpoints will be at the two-sided 95% level. All p-values reported will be two-sided. P-values will be rounded to the fourth decimal place. Point estimates and confidence bounds for efficacy variables will be rounded to the second decimal place.

A by-subject listing of efficacy results will be presented including treatment group, treatment duration, BICR progression date, overall survival, death date, etc.

7.5.1 Analysis of Progression-Free Survival

The primary objective of the study is to compare the PFS per BICR of Arm A to Arm C in randomized subjects with previously untreated (first line) advanced or metastatic RCC. All the analyses outlined in this section are specified for the all randomized subjects population in Arm A and Arm C, unless otherwise specified.

PFS per BICR will be compared between the treatment groups via stratified log-rank test among all randomized subjects at a two-sided $\alpha = 0.05$ level. The stratification factors will be IMDC prognostic risk score (0 vs 1-2 vs 3-6), region (US/Canada/W Europe/N Europe vs ROW)' and PD-L1 status (≥ 1% vs < 1% or indeterminate).

The primary definition of PFS adjusting for subsequent anticancer therapy will be used in this analysis. The two-sided log-rank p-value will be reported.

The estimate of the PFS hazard ratio between treatment groups will be calculated using a stratified Cox proportional hazards model, with treatment as the sole covariate. Ties will be handled using the exact method. A two-sided 95% CI for the hazard ratio will also be presented.

The PFS function for each treatment group will be estimated using the KM product limit method and will be displayed graphically. A two-sided 95% CI for median PFS in each treatment group will be computed via the log-log transformation method. PFS rates at fixed time points (e.g. 6 months, depending on the minimum follow-up) will be presented along with their associated 95% CIs. These estimates will be derived from the Kaplan Meier estimate and corresponding CIs will be derived based on Greenwood[11] formula for variance derivation and on log-log transformation applied on the survivor function[12].

Analyses of PFS will also be conducted based on the secondary definition of PFS. These analyses will be the same as those specified above.

The source of PFS event (progression or death) will be summarized by treatment group. The status of subjects who are censored (as per primary definition of PFS) in the PFS KM analysis will be tabulated for each treatment group including the following categories:

- On-study (on-treatment, in follow-up)
- Off-study (lost to follow-up, withdraw consent, never treated)
- No baseline tumor assessment
- No on-study tumor assessment and no death
- Received subsequent anticancer therapy

A by-subject listing will be presented including treatment group, PFS duration under the primary definition, PFS duration on the ITT definition, whether the subject was censored under the primary definition, and if censored, the reason, and whether the subject was censored under the ITT definition, and if censored, the reason.

A by-subject listing of lesion evaluations per BICR will be presented.

Note that similar tables (KM plot, median PFS with its 95% CI, PFS rates, source of PFS event, and status of subjects who are censored) along with by-subject listing will be summarized for all intermediate/poor risk subjects in Arm A and Arm C.

For those subjects who randomized to Arm B prior to Revised Protocol 01, the PFS function will be estimated using the KM product limit method and reported seperately for Arm B. A two-sided

95% CI for median PFS will be computed via the log-log transformation method. A by-subject listing will be presented including treatment group, PFS duration under the primary definition, PFS duration on the ITT definition, whether the subject was censored under the primary definition, and if censored, the reason, and whether the subject was censored under the ITT definition, and if censored, the reason.

7.5.2 Supportive Analyses of Progression-Free Survival

The following sensitivity analyses will be conducted using both the primary and the secondary definition of PFS in all randomized subjects:

1) Delayed effect of immunotherapy interventions may cause a late separation in the PFS KM curves and non-proportional hazards. PFS (as determined by BICR) will be compared between treatment groups via two-sided 0.05 stratified weighted log-rank test among subjects. The primary definition of PFS will be used in this analysis. The two-sided stratified weighted log-rank p-value will be reported using G (rho = 0, gamma = 1) weights, in the terminology of Fleming and Harrington[13].
   a) The Fleming Harrington test can be unstable, so it is possible, though uncommon, that thep-value for this trial will not be estimable.
   b) The estimate of the PFS hazard ratio in the period before and following 6 months will be calculated using a stratified time-dependent Cox model with effects for treatment and period-by-treatment interaction. In this model, period is a binary variable indicating pre- vs. post-6 months. Ties will be handled using the exact method. A two-sided 95% CI for the hazard ratio will also be presented.
2) A multivariate Cox regression model will be used in order to estimate the treatment effect after adjustment for possible imbalances in known or potential prognostic factors. The factors used in the randomization, which, by definition, will be balanced across treatment groups, will still be included in the model as stratification factors. However, all additional factors will be incorporated as covariates. The additional factors, which are all measured at baseline, will include:
   a) Age categorization (< 65 vs. ≥ 65)
   b) Gender (Male vs. Female)
   c) Race
   d) Region (US/Canada/W.Europe/N.Europe vs. ROW)
   e) IMDC score (0 vs 1-2 vs 3-6)
   f) Karnofsky performance status (100-90, <90)
   g) Prior Nephrectomy (Yes, No)
   h) LDH level (≤ 1.5 x ULN , > 1.5 x ULN)
   i) Baseline PD-L1+ status based on a 1% cut off
   j) Number of organ with metastasis (1 vs. ≥2)

   The level of the covariate normally associated with the worst prognosis will be coded as the reference level. The hazard ratio associated with treatment and with each of the baseline covariates will be presented along with associated 95% CIs.

3) PFS using stratification factors as obtained from the baseline CRF pages (instead of IRT). The hazard ratio associated with treatment will be presented along with the associated two-sided 95% CIs. This analysis will be performed only if at least one stratification variable/factor at randomization (as per IRT) and baseline are not concordant for at least 10% of the randomized subjects.
4) PFS using the investigator's assessment. The hazard ratio associated with treatment and median PFS will be presented along with the associated two-sided 95% CIs.

   A cross tabulation of PFS assessment by BICR versus PFS assessment by investigator will be presented, by treatment group. Concordance Rate of event will be computed as the frequency with which investigator and BICR agree on classification of a subject as event versus censored as a proportion of the total number of randomized subjects assessed by both the investigator and BICR.

   A by-subject listing of PFS assessment per BICR and investigator will be presented.
5) PFS using an un-stratified log rank test. The hazard ratio associated with treatment will be presented along with the associated two-sided 95% CIs.
6) PFS using an un-stratified Cox proportional hazards model, adjusted, using as covariates only the stratification factors used in randomization. The hazard ratio associated with treatment will be presented along with the associated two-sided 95% CIs.
7) PFS for subjects with no relevant protocol deviations. This analysis will be conducted only if there are more than 10% subjects with relevant protocol deviations. The hazard ratio associated with treatment will be presented along with the associated two-sided 95% CIs.
8) The method of Gail and Simon14 will be used to test for a qualitative interaction between treatment and strata. This test will be conducted at $\alpha = 0.10$ level. The p-value reported from this specific analysis is for descriptive purposes alone.
9) To examine the assumption of proportional hazards in the Cox regression model, in addition to treatment, a time-dependent variable defined by treatment by time interaction will be added into the model. A two-sided Wald Chi-square p-value of less than 0.1 may indicate a potential nonconstant treatment effect. In that case, additional exploratory analyses may be performed.
10) PFS using censoring for 2 missing TA

7.5.3 Subset Analyses of Progression-Free Survival

The influence of baseline and demographic characteristics on the treatment effect among all randomized subjects will be explored via exploratory subset analyses. The median PFS based on KM product-limit method along with two-sided 95% CIs will be produced for the following subgroups:

- Age
- Age categorization (< 65, ≥ 65 and < 75, ≥ 75 and < 85, ≥ 85, ≥ 75, ≥ 65)
- Sex (Male, Female)
- Race
- Region (Region (US/Canada/W.Europe/N.Europe vs. ROW) (source: IRT)

- Ethnicity (Hispanic/Latino and Not Hispanic/Latino)
- Karnofsky performance status (100-90, <90)
- Baseline IMDC prognostic score (0, 1-2, ≥ 3) (source: IRT)
- Baseline IMDC prognostic score (0, 1-2, ≥ 3) (source: CRF)
- Time from initial disease diagnosis to randomization (<1 year, ≥1 year)
- Baseline LDH level (≤ 1.5 x ULN, >1.5 x ULN)
- Hemoglobin (<LLN, ≥ LLN)
- Corrected Calcium (≤ 10 mg/dl, >10mg/dl)
- Absolute Neutrophil Count (≤ ULN, > ULN)
- Platelet Count (≤ ULN, > ULN)
- Baseline Alkaline phosphatase (< ULN, ≥ ULN)
- Prior nephrectomy (Yes, No)
- Prior radiotherapy (Yes, No)
- Bone metastasis (Yes, No)
- Sarcomatoid features (Yes, No)
- Stage at the initial diagnosis (Stage IV, non-Stage IV)
- Baseline PD-L1+ status based on a 1% cut off
- Baseline PD-L1+ status based on a 5% cut off
- Baseline PD-L1+ status based on a 10% cut off

A forest plot of the unstratified PFS hazard ratios (along with the 95% CIs) will be produced for each level of the subgroups listed above. The analysis comparing treatment (i.e., Hazard Ratio) will be conducted if the number of subjects in the subgroup category is more than 10.

7.5.4 Analysis of Overall Survival

One of the secondary objectives of the study is to compare the overall survival of Arm A to Arm C in all randomized subjects. All the analyses outlined in this section are specified for the all randomized subjects population in Arm A and Arm C, unless otherwise specified.

If the formal analysis of PFS among all randomized subjects is statistically significant, the formal interim analysis of OS among all randomized subjects will be tested, as per hierarchical testing procedure.

Overall survival will be compared between the treatment groups at the interim and final analyses, using stratified log-rank test. The stratification factors will be those used in the analysis of PFS. An O'Brien and Fleming α-spending function will be employed to determine the nominal significance levels for the interim and final analyses. The stratified hazard ratio between the treatment groups will be presented along with 100*(1- α)% CI (adjusted for interim). In addition, two-sided p-value will also be reported for the analysis of OS.

OS will be estimated using the KM techniques. A two-sided 95% CI for median OS in each treatment group will be computed via the log-log transformation method. OS rates at fixed time points (e.g. 6 months, depending on the minimum follow-up) will be presented along with their associated 95% CIs. These estimates will be derived from the Kaplan Meier estimate and corresponding CIs will be derived based on Greenwood formula for variance derivation and on log-log transformation applied on the survivor function.

The status of subjects who are censored in the OS KM analysis will be tabulated for each treatment group using the following categories:

- On-study (on-treatment, in follow-up)
- Off-study (lost to follow-up, withdraw consent, never treated)

A by-subject listing will be presented including treatment group, first and last dose date, whether the subject died, and if censored, the reason, event/censored date and OS duration.

Note that similar tables (KM plot, median OS with its 95% CI, OS rates, and status of subjects who are censored) along with by-subject listing will be summarized for all intermediate/poor risk subjects in Arm A and Arm C.

The analysis performed for PFS (detailed in section 7.5.8) for those subjects who randomized to Arm B prior to Revised Protocol 01 will be repeated for OS.

7.5.5 Subset Analysis of Overall Survival

The influence of baseline and demographic characteristics on the treatment effect among all randomized subjects will be explored via exploratory subset analyses. The median OS based on KM product-limit method along with two-sided 95% CIs will be produced for the same subgroups as used for PFS (see Section APPENDIX 1).

A forest plot of the unstratified OS hazard ratios (along with the 95% CIs) will be produced for each level of the subgroups listed above.

An analysis will be conducted if the number of subjects in the subgroup category is more than 10.

7.5.6 Current Status of PFS and OS Follow-up

The extent of follow-up for survival, defined as the time between randomization date and last known alive date (for subjects who are alive) or death date (for subjects who died), will be summarized descriptively (median, min, max, etc.) in months for all randomized subjects.

The currentness of follow-up for survival, defined as the time between last OS contact (i.e., last known alive date or death date) and cutoff date (defined by last subject last visit date), will be summarized in months for all randomized subjects. Subjects who died and subjects with last known alive date on or after data cut-off date will have zero value for currentness of follow-up.

Minimum follow-up of OS for all randomized subjects, defined as the time from cutoff date to last subject's randomization date, will be summarized in months.

Time from last evaluable tumor assessment to cutoff date in months will be summarized by treatment group and overall for all randomized subjects. Subjects who have a PFS event will be considered as current for this analysis. The secondary definition of PFS will be used for this summary.

In addition, time to treatment discontinuation will be summarized and presented by treatment group using a Kaplan-Meier curve whereby the last dose date will be the event date for those subjects who are off study therapy. Median duration of study therapy and associated 95% CI will be provided. Subjects who are still on study therapy will be censored on their last dose date.

A by-subject listing will also be produced to accompany the subject time from last evaluable tumor assessment.

7.5.7 Interim Analysis of Overall Survival

An independent statistician external to BMS will perform the analysis. In addition to the formal planned interim analyses for OS, the Data Monitoring Committee (DMC) will have access to periodic un-blinded interim reports of efficacy and safety to allow a risk/benefit assessment. Details are included in the DMC charter.

Two interim analyses of OS are planned for this study. The first interim analysis of OS is planned at the time of final PFS analysis and expected after observing 165 deaths (approximately 65% of the targeted OS events) have been observed among all randomized subjects in Arm A and Arm C based on above accrual rate and the exponential distribution in each arm. These formal comparisons of OS will allow for early stopping for superiority, and the boundaries for declaring superiority will be derived based on the actual number of deaths using Lan-DeMets spending function with O'Brien and Fleming type of boundary in EAST version 6. If the first interim analysis is performed exactly at 165 deaths, the boundary in terms of statistical significance for declaring superiority would be 0.011 (HR=0.673 with 16 months improvement in median OS for the Arm A versus Arm C comparison (33 versus 49 months)). The second interim analysis of OS is expected after observing 211 deaths (approximately 83% of the targeted OS events) have been observed among all randomized subjects based on above accrual rate and the exponential distribution in each arm. The boundary for declaring superiority in terms of statistical significance for the second interim analysis after 211 events would be 0.025 (HR=0.734 with 12 months improvement in median OS for the doublet versus sunitinib comparison (33 versus 45 months). The boundary for declaring superiority in terms of statistical significance for the final analysis after 254 events would be 0.041 (HR=0.774 with 9.6 months improvement in median OS for the Arm A versus Arm C comparison (33 versus 42.6 months).

Note that if the analysis of PFS final analysis and first interim analysis OS is trigger with 8-months minimum follow-up on all randomized subjects, minimum 283 PFS events and 149 OS events, then the details of the first interim analysis of OS will be as follows:

- If the first interim analysis is performed exactly at 149 deaths, the boundary in terms of statistical significance for declaring superiority would be 0.007 (HR=0.643 with 18.3 months improvement in median OS for the Arm A versus Arm C comparison (33 versus 51.3 months)).

The DMC will review the safety and efficacy data from the informal interim analyses, BMS will remain blinded to these interim results and and will determine if the study should continue with or without changes or if accrual should be stopped. Subject enrollment will continue while waiting for the DMC's decisions.

The chair of the DMC and the sponsor can call an unscheduled review of the safety data.

If the formal analysis of PFS among all randomized subjects is statistically significant, the formal interim analysis of OS among all randomized subjects will be tested, as per hierarchical testing procedure.

At the time of the formal interim analysis for superiority of OS, the DMC may recommend continuing or stopping the trial. If the trial continues beyond the formal interim analysis, BMS will remain blinded to these interim results and the nominal critical point for the final OS analysis will be determined using the recalculated information fraction at the time of the interim analysis, as described above. The final OS hazard ratio and corresponding confidence interval will be reported whereby the confidence interval will be adjusted accordingly (i.e. using the recalculated nominal α level at the final analysis).

If the trial is stopped for superiority of OS at the interim, the p-value from the interim stratified log-rank test will be considered the final primary analysis result.

7.5.8 Analysis of Objective Response Rate

One of the secondary objectives of the study is to evaluate the objective response rate in all randomized subjects in Arm A and Arm C. All the analyses outlined in this section are specified for the all randomized subjects population in Arm A and Arm C, unless otherwise specified. If the formal analysis of OS among all randomized subjects is statistically significant, the formal analysis of ORR among all randomized subjects will be tested, as per hierarchical testing procedure. Note that the formal ORR analysis specified in the SAP supersede those specified in the protocol.

The number and percentage of subjects in each category of BOR per BICR (complete response [CR], partial response [PR], stable disease [SD], progressive disease [PD], or unable to determine [UTD]) will be presented, by treatment group. Estimates of response rate, along with its exact two-sided 95% CI by Clopper and Pearson[15] will be presented, by treatment group.

Similar analyses will be repeated based on the investigator's assessment of ORR. A cross tabulation of BICR best response versus the investigator best response will be presented, by treatment group and by response categories. Concordance Rate of Responders will be computed as the frequency with which investigator and BICR agree on classification of a subject as responder vs. non responder/UTD as a proportion of the total number of randomized subjects assessed by both the investigator and BICR.

The following subject-level graphics will also be provided:

- For the responders only, time courses of the following events of interest will be graphically displayed: tumor response, progression, last dose received, and death.

- For response evaluable subjects (randomized subjects with baseline and at least one on-study tumor assessment),
  - A bar plot showing the best % reduction from baseline in sum of diameter of target lesions based on BICR assessment for each subject will be produced (excluding assessments after PD and assessments after start of subsequent anti-cancer therapy).
  - A plot of individual time course of tumor burden change per BICR assessment will be produced.

A by-subject listing of best overall response will be presented including treatment group, best overall response per BICR and dates of CR/PR/progression.

A by-subject listing of per time point tumor response per BICR will be presented.

Note that similar tables along with by-subject listing will be summarized for all intermediate/poor risk subjects in Arm A and Arm C.

For those subjects who randomized to Arm B prior to Revised Protocol 01, estimates of response rate, along with its exact two-sided 95% CI by Clopper-Pearson method, will be computed per BICR and investigator. DOR and TTR will also be evaluated per BICR.

7.5.9 Subset Analyses of Objective Response

The influence of baseline and demographic characteristics on the treatment effect will be explored via exploratory subset analysis. The subsets wil be same as those analyzed for PFS and will be reported based on the BICR assessment of ORR.

A forest plot of treatment effect on ORR per BICR in the above subgroups will be produced. The un-weighted differences in ORR between the two treatment groups and corresponding 95% two-sided CI using the method of Newcombe will be provided if the number of subjects in the subgroup category is more than 10.

7.5.10 Time to Tumor Response and Duration of Response

The analyses specified in this section will be conducted for all treatment arms. Duration of response (DOR) and time to response (TTR) will also be evaluated for subjects who achieved confirmed PR or CR. The DOR for each treatment group will be estimated using the Kaplan-Meier (KM) product limit method and will be displayed graphically. A table will be produced presenting number of events, number of subjects involved, medians, and 95% CIs for the medians. Median values of DOR, along with two-sided 95% CI in each treatment group will be computed based on a log-log transformation method.

The status of subjects who are censored in the DOR KM analysis will be tabulated for each treatment group including the following categories:

- Ongoing follow-up (current [last scan within adequate window vs cutoff date], not current)
- Off-study (lost to follow-up, withdraw consent, never treated)
- Received subsequent anticancer therapy.

TTR, which does not involve censoring, will be summarized by treatment group in all responders using descriptive statistics.

Cumulative Response Rates will be tabulated for Week 8, Month 4, 6, 8, and 12, and overall response rate will be provided.

A by-subject listing will be presented including treatment group, best response, time to response, duration of response, whether the subject was censored for duration of response, and, if so, the reason.

7.5.11 PFS2

One of the exploratory objectives of the study is to compare PFS2 between treatment groups in all randomized subjects.

PFS2 will be analyzed similarly to PFS:

- Median values based on KM method, along with two-sided 95% CI using Brookmeyer and Crowley method will be calculated. The estimate of standard error will be calculated using the Greenwood formula;
- PFS2 will be graphically displayed along with the median and 95% CI.

A by-subject listing of PFS and PFS2 will be provided.

7.6 Safety

Analyses in this section will be tabulated for all treated subjects by treatment group as treated, unless otherwise specified.

Analyses in this section will be tabulated for all treated subjects in Arm A and Arm C, unless otherwise specified. Limited selection of the summary tables will be generated for those subjects who randomized to Arm B prior to Revised Protocol 01, which is specified in the relevant subsections below.

7.6.1 Deaths

Deaths will be summarized by treatment group:

- All deaths, reasons for death.
- Deaths within 30 days of last dose received, reasons for death.
- Deaths within 100 days of last dose received, reasons for death.

A by-subject listing of deaths will be provided for the all enrolled subjects population.

Note that for all intermediate/poor risk treated subjects in Arm A and Arm C, deaths will be summarized by treatment group.

Similar tables will be presented for those subjects treated in Arm B.

7.6.2 Serious Adverse Events

Serious adverse events will be summarized by treatment group:

- Overall summary of SAEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.
- Overall summary of drug-related SAEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.

All analyses will be conducted using the 30-day safety window.

A by-subject SAE listing will be provided for the "enrolled subjects" population.

Note that for all intermediate/poor risk treated subjects in Arm A and Arm C, serious adverse events will be summarized by treatment group.

Similar tables will be presented for those subjects treated in Arm B.

7.6.3 Adverse Events Leading to Discontinuation of Study Therapy

AEs leading to discontinuation will be summarized by treatment group:

- Overall summary of AEs leading to discontinuation by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.
- Overall summary of drug-related AEs leading to discontinuation by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.

The analyses will be conducted using the 30-day safety window.

A by-subject AEs leading to discontinuation listing will be provided.

Note that for all intermediate/poor risk treated subjects in Arm A and Arm C, AEs leading to discontinuation will be summarized by treatment group.

Similar table will be presented for those subjects treated in Arm B.

Note that

- for Arm A, AEs leading to discontinuation from cabozantinib only, nivolumab only, and both cabozantinib and nivolumab.
- for Arm B, AEs leading to discontinuation from cabozantinib only, nivolumab and ipilimumab only, and from cabozantinib, nivolumab, and ipilumumab will be summarized in all the tables specified in this section.

7.6.4 Adverse Events Leading to Dose Modification

AEs leading to dose delay/reduction will be summarized by treatment group:

- Overall summary of AEs leading to dose delay/reduction by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.

- Overall summary of related AEs leading to dose delay/reduction by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.

The analysis will be conducted using the 30-day safety window.

A by-subject AEs leading to dose delay/reduction listing will be provided.

7.6.5 Adverse Events

Adverse events will be summarized by treatment group.

The following analyses will be conducted using the 30 days safety window only:

- Overall summary of any AEs by worst CTC grade (1, 2, 3, 4, 5, not reported, total) presented by SOC/PT.
- Overall summary of any AEs presented by worst CTC grade (any grade, grade 3-4, grade 5) by SOC/PT. This table will be restricted to events with an incidence greater or equal to 5% in any treatment group.
- Overall summary of any non-serious AEs presented by SOC/PT. This table will be restricted to events with an incidence greater or equal to 5% in any treatment group.
- Overall summary of any AEs that required immune modulating medication by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.
- Overall summary of drug-related AEs by worst CTC grade (1, 2, 3, 4, 5, not reported, total) presented by SOC/PT.

The following analyses will be conducted using the 30 days safety window and repeated using the 100 days safety window:

- Overall summary of drug-related AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.

A by-subject AE listing will be provided. A by-subject listing of any AE requiring immune modulating medications will also be provided.

For those subjects treated in Arm B and for those intermediate/poor risk treated subjects in Arm A and Arm C:

The following analyses will be conducted using the 30 days safety window only:

- Overall summary of any AEs by worst CTC grade (1, 2, 3, 4, 5, not reported, total) presented by SOC/PT.
- Overall summary of any AEs that required immune modulating medication by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.
- Overall summary of drug-related AEs by worst CTC grade (1, 2, 3, 4, 5, not reported, total) presented by SOC/PT.

The following analyses will be conducted using the 30 days safety window and repeated using the 100 days safety window:

- Overall summary of drug-related AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.

A by-subject AE listing will be provided. A by-subject listing of any AE requiring immune modulating medications will also be provided

7.6.6 Select Adverse Events (EU/ROW Submissions)

Unless otherwise specified, analyses will be performed by select AE category. Analyses will also be repeated by subcategory of endocrine events.

7.6.6.1 Incidence of Select AE

Select AEs will be summarized by treatment group for each category/subcategory.

The following analyses will be conducted using the 30-day safety window only:

- Overall summaries of any select AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category or Subcategory/PT.
- Overall summaries of any drug-related select AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category or Subcategory/PT.
- Overall summaries of any serious select AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category or Subcategory /PT.
- Overall summaries of drug-related serious select AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category or Subcategory /PT.
- Overall summaries of any select AEs leading to discontinuation by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category or Subcategory /PT.
- Overall summaries of drug-related select AEs leading to discontinuation by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category or Subcategory /PT.
- Summary of frequency of unique select AEs by Category.

A by-subject select AE listing will be provided.

7.6.6.2 Time-to Onset of Select AE

Time-to onset of drug-related select AEs (any grade, grade 3-5) will be summarized for each category/subcategory by treatment group.

Time-to onset analyses are restricted to treated subjects who experienced at least one drug-related select AE in the category/subcategory. The analyses will be conducted using the 30-day safety window.

Additional details regarding the time-to onset definition are described in time-to onset definition subsection of APPENDIX 1.

7.6.6.3 Time-to Resolution of Select AE

Time-to resolution of the following specific events will be summarized separately for each category/subcategory.

- Time-to resolution of drug-related select AE (any grade, grade 3-5) by treatment group
- Time-to resolution of drug-related select AE (any grade, grade 3-5) where immune modulating medication was initiated, by treatment group

Time-to resolution analyses are restricted to treated subjects who experienced the specific events. Time-to resolution where immune modulating medication was initiated analyses are restricted to treated subjects who experienced the specific events and who received immune modulating medication during the longest select AE.

The analyses will be conducted using the 30-day safety window.

The following summary statistics will be reported: percentage of subjects with resolution of the longest select AE, median time-to resolution along with 95% CI (derived from Kaplan-Meier estimation) and ranges.

See time-to resolution definition subsection of APPENDIX 1 for additional details.

7.6.7 Immune-Mediated Adverse Events (US Submission)

IMAEs will be summarized by treatment group for each immune-mediated category / PT using the 100-day safety window:

- Overall summary of non-endocrine IMAEs by worst CTC grade (any grade, grade 3-4, grade 5) where immune modulating medication was initiated presented by Category / PT.
- Overall summary of endocrine IMAEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category / PT.
- Overall summary of non-endocrine IMAEs leading to discontinuation by worst CTC grade (any grade, grade 3-4, grade 5) where immune modulating medication was initiated presented by Category / PT.
- Overall summary of endocrine IMAEs leading to discontinuation by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category / PT.
- Overall summary of non-endocrine IMAEs leading to dose delay or reduction by worst CTC grade (any grade, grade 3-4, grade 5) where immune modulating medication was initiated presented by Category / PT
- Overall summary of endocrine IMAEs leading to dose delay or reduction by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category / PT.
- Summaries of time-to onset and time-to resolution of non-endocrine IMAEs where immune modulating medication was initiated presented by Category.
- Summaries of time-to onset and time-to resolution of endocrine IMAEs presented by Category.

A by-subject listing of IMAEs will be provided. By-subject listings of time-to resolution for longest IMAEs cluster (any grade and grade 3-5 in separate summaries) will also be provided. For new studies which collect investigator assessment of potential IMAE data, a by-subject listing of AEs considered as immune-mediated events per investigator but not qualified for IMAEs definition will also be provided.

In addition, for all nivolumab treated subjects who experienced at least one IMAE, the following data presentation will be provided:

- Summary of subjects who were re-challenged with nivolumab by IMAE category, with extended follow-up
- Summary of subjects who were re-challenged with nivolumab or ipilimumab by IMAE category, with extended follow-up

For these, re-challenge is considered to have occurred when last nivolumab and/or ipilimumab infusion was administered after the onset of an IMAE.

7.6.8 Other Events of Special Interest

OEOSI will be summarized by treatment group for each category.

The following analyses will be conducted using the 100-day safety window:

- Overall summary of OEOSI by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category / PT
- Overall summary of drug-related OEOSI by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category / PT

A by-subject listing of OEOSI will be provided.

7.6.9 Multiple Events

The following summary tables will be provided:

- A table showing the total number and rate (exposure adjusted) of occurrences for all AEs.
- A table showing the total number and rate (exposure adjusted) of occurrences for AEs occurring in at least 5% of subjects in any treatment group.

In addition, the rate (exposure adjusted) and its 95% CI evaluated for different time intervals will be displayed graphically for each treatment group. This analysis will be limited to the rate of all AEs and all drug-related AEs. The analyses will be conducted using the 30-day safety window.

A listing displaying the unique instances of all AEs, i.e., after duplicates have been eliminated and overlapping and contiguous occurrences of the same event (i.e. same PT) have been collapsed will be provided. No formal comparisons will be made between treatment groups.

7.6.10 Laboratory Parameters

The analysis population for each laboratory test is restricted to treated subjects who underwent that laboratory test. Laboratory tests (in addition to the tests specified below) with CTC criteria collected in the specific studies may also be included in the summaries.

A by-subject listing of differences in categorization of SI and US laboratory test results will be provided.

7.6.10.1 Hematology

The following will be summarized by treatment group as worst CTC grade on-treatment per subject and as shift table of worst on-treatment CTC grade compared to baseline CTC grade per subject: hemoglobin (HB), platelets, white blood counts (WBC), absolute neutrophils count (ANC) and lymphocyte count (LYMPH).

The analyses will be conducted using the 30-day safety window.

A by-subject listing of these laboratory parameters will be provided.

7.6.10.2 Serum Chemistry

The following will be summarized by treatment group as worst CTC grade on-treatment per subject and as shift table of worst on-treatment CTC grade compared to baseline CTC grade per subject: ALT, AST, alkaline phosphatase (ALP), total bilirubin and creatinine.

The analyses will be conducted using the 30-day safety window.

A by-subject listing of these laboratory parameters will be provided.

7.6.10.3 Electrolytes

The following will be summarized by treatment group as worst CTC grade on-treatment per subject and as shift table of worst on-treatment CTC grade compared to baseline CTC grade per subject: sodium (high and low), potassium (high and low), calcium (high and low), magnesium (high and low), and Glucose Serum (fasting hyperglycemia and hypoglycemia regardless of fasting status).

The analyses will be conducted using the 30-day safety window.

A by-subject listing of these laboratory parameters will be provided.

7.6.10.4 Additional Analyses

In addition, further analyses on specific laboratory parameters will be performed by treatment group:

Abnormal Hepatic Function Test

The number of subjects with the following laboratory abnormalities from on-treatment evaluations will be summarized by treatment group:

- ALT or AST > 3 x ULN, > 5 x ULN, > 10 x ULN and > 20 x ULN
- Total bilirubin > 2 x ULN

- ALP > 1.5 x ULN
- Concurrent (within 1 day) ALT or AST > 3 x ULN and total bilirubin > 1.5 x ULN
- Concurrent (within 30 days) ALT or AST > 3 x ULN and total bilirubin > 1.5 x ULN
- Concurrent (within 1 day) ALT or AST > 3 x ULN and total bilirubin > 2 x ULN
- Concurrent (within 30 days) ALT or AST > 3 x ULN and total bilirubin > 2 x ULN

The analyses will be conducted using the 30-day safety window.

A by-subject listing of these specific abnormalities will be provided.

Abnormal Thyroid Function Test

The number of subjects with the following laboratory abnormalities from on-treatment evaluations will be summarized by treatment group:

- TSH value > ULN and
  - with baseline TSH value ≤ ULN
  - with at least one FT3/FT4 test value < LLN within 2-week window after the abnormal TSH test
  - with all FT3/FT4 test values ≥ LLN within 2-week window after the abnormal TSH test
  - with FT3/FT4 missing within 2-week window after the abnormal TSH test.
- TSH < LLN and
  - with baseline TSH value ≥ LLN
  - with at least one FT3/FT4 test value > ULN within 2-week window after the abnormal TSH test
  - with all FT3/FT4 test values ≤ ULN within 2-week window after the abnormal TSH test
  - with FT3/FT4 missing within 2-week window after the abnormal TSH test

The analyses will be conducted using the 30-day safety window.

A by-subject listing of these specific abnormalities will be provided.

7.6.11 Vital Signs and Pulse Oximetry

Vital signs and pulse oximetry (i.e. % oxygen saturation) collected on the CRF will be provided in separate listings.

7.6.12 Physical Measurements

Physical measurements will be listed by subject.

7.6.13 Non-Protocol Medical Procedures

Non-protocol medical procedures will be listed by subject.

7.6.14 Immunogenicity Analysis

Further details on immunogenicity background and rationale, definitions, population for analyses and endpoints are described in APPENDIX 3.

Incidence of ADA

Number (%) of subjects will be reported for the following parameters based on Evaluable Subjects.

- Baseline ADA-positive
- ADA-positive
  - Persistent Positive (PP)
  - Not PP-Last Sample Positive
  - Other positive
  - Neutralizing Positive
- ADA-negative

A listing of all ADA assessments will be provided. A separate listing of ADA assessments for subjects with neutralizing positive will also be provided.

A spider plot of nivolumab ADA test result (titers) over time may be provided for nivolumab ADA positive subjects.

Clinical implications

Clinical implications of nivolumab immunogenicity will be primarily focused on subjects with persistent ADA-positive relative to ADA-negative. Subjects with any ADA-positive samples after initiation of treatment (relative to baseline) may be used to explore clinical implications. Effect of immunogenicity on clearance of nivolumab will be explored by comparison of clearance estimates (determined by PPK analysis). Effect of immunogenicity on safety will be explored by examining the frequency and type of AEs of special interest such as hypersensitivity/infusion reaction. Summary tables for incidence of overall and each of the preferred terms will be provided, if the number of subjects is of sufficient size (e.g., at least 10 subjects). Otherwise, individual subject's safety profile will be examined and described based on a listing. Clinical implications on efficacy will also be explored similarly. Association between trough concentrations of nivolumab or combination drug (e.g., ipilimumab) and ADA assessments may be explored, as needed.

The following data presentation will be provided:

- Swimmer plot of occurrence of ADA and NAb Occurrence in Relation to PFS, BOR per investigator and OS

7.6.15 Pregnancy

A by-subject listing of pregnancy tests results will be provided for randomized female subjects.

7.6.16 Adverse Events By Subgroup

Overall summary of any AEs and drug-related AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT and for each treatment group for the following subgroups:

- Sex (Male vs. Female)
- Race
- Age (< 65 vs. 65 - < 75 vs. 75 - < 85 vs. ≥ 85 vs. ≥ 75 vs. ≥ 65)
- Region (US/Canada/W.Europe/N.Europe vs. ROW)

These analyses will be conducted using the 30-day safety window only.

7.7 Pharmacokinetics

The nivolumab and cabozantinib concentration data obtained in this study will be combined with data from other studies in the clinical development program to develop a population PK model. This model will be used to evaluate the effects of intrinsic and extrinsic covariates on the PK of nivolumab and cabozantinib. In addition, exposure-response analyses with selected efficacy and safety endpoints will be conducted. Results of population PK and exposure response-analyses will be reported separately.

7.8 Biomarkers

Analyses for PD-L1 are described below.

7.8.1 Distribution of PD-L1 Expression

Descriptive statistics of PD-L1 expression:

- Listing of all PD-L1 IHC data, all randomized subjects
- Summary of tumor specimen acquisition and characteristics, all randomized subjects
- Summary statistics of PD-L1 expression in all randomized subjects with quantifiable PD-L1 expression
- Summary of BOR and ORR by baseline PD-L1 expression at 1, 5, and 10% cutoffs in all randomized subjects
- Kaplan-Meier curves for PFS and OS by BICR by baseline expression of PD-L1 < 1%, >= 1 and < 5, > = 5 and < 10, and >= 10
- Cumulative distribution plot of baseline PD-L1 expression versus population percentile in all randomized subjects with quantifiable PD-L1 expression
- Box plots of PD-L1 expression at 1, 5, and 10% cutoffs versus Response Status (BICR assessment) in all randomized subjects with quantifiable PD-L1 expression
- Waterfall plot of individual PD-L1 expression in all randomized subjects with quantifiable PD-L1 expression

7.8.2 Other Exploratory Biomarkers

The following analyses for other exploratory biomarkers will be provided when the data becomes available, which may occur after the CSR is written and therefore a seperate Data Presentation Plan (DPP) will be developed:

Descriptive statistics of MDSC expression at baseline:

- Listing of all MDSC data, all randomized subjects
- Summary statistics of MDSC expression in all randomized subjects with quantifiable MDSC expression
- Correlation statistics of BOR and ORR by baseline MDSC expression levels
- Cumulative distribution plot of baseline MDSC expression versus population percentile in all randomized subjects with quantifiable MDSC expression
- Kaplan-Meier curves for PFS and OS by BICR by baseline expression of tertitle MDSC expression
- Box plots of MDSC expression at tertile cutoffs versus Response Status (BICR assessment) in all randomized subjects with quantifiable MDSC expression

Descriptive statistics of c-MET expression at baseline:

Summary statistics of c-MET expression in all randomized subjects will be reported.

Descriptive statistics of serum cytokine expressions (only MIG, IP10 and IFN-g) changes in the course of treatment:

- Line plot per arm for each cytokine on X axis and with timeline on Y axis with statistical difference from the baseline to be marked

Descriptive statistics for gene expression signature analyses:

- Published, on-target and mechanism of action related pathway related gene expression signatures (e.g. CD8, angiogenesis etc) will be assessed in BOR, ORR, PFS and OS relationship

7.9 Clinical Outcomes Assessments

The analysis of FKSI-19 and EQ-5D-3L will be restricted to randomized subjects in Arm A and C who have an assessment at baseline and at least one post-baseline assessment.

7.9.1 FKSI-19

Unless otherwise specified, the analysis of FKSI-19 will be performed by treatment group in all randomized subjects who have an assessment at baseline and at least one or more post-baseline assessments. The following descriptive analyses will be conducted:

- Questionnaire completion rate, defined as the proportion of questionnaires actually received out of the expected number (i.e., number of subjects on treatment or in follow up), will be calculated and summarized for each assessment time point.
- For the total score and subscales, separately:
  - Mean score and mean change from baseline at each assessment time point will be summarized using descriptive statistics (N, mean, SD, median, 25th and 75th percentiles, minimum, maximum).
  - A plot summarizing the mean change from baseline will be presented and including 95% CI.

By subject listing of FKSI-19 will be provided.

7.9.2 EuroQol EQ-5D-3L

The following descriptive analyses will be conducted:

- EQ-5D-3L questionnaire completion rate, defined as the proportion of questionnaires actually received out of the expected number (i.e. number of subjects on treatment or in follow up), will be calculated and summarized for each assessment time point by treatment group.
- A by-subject listing of the level of problems in each dimension, corresponding to EQ-5D-3L health state (i.e., 5-digit vector), EQ-5D-3L utility index score, and EQ-5D-3L VAS score will be provided.
- Proportion of subjects reporting problems for the 5 EQ-5D-3L dimensions at each assessment time point will be summarized by level of problem and by treatment group. Percentages will be based on number of subjects assessed at assessment time point.
- For the EQ-5D-3L utility index and VAS scores, separately:
  - Mean score and mean change from baseline at each assessment time point will be summarized by treatment group using descriptive statistics (N, mean with SD and 95% CI, median, first and third quartiles, minimum, maximum).
  - A line graph summarizing the mean changes from baseline will be produced.

8 CONVENTIONS

The following conventions may be used for imputing partial dates for analyses requiring dates:

- For missing and partial adverse event onset dates, imputation will be performed using the Adverse Event Domain Requirements Specification[16]
- For missing and partial adverse event resolution dates, imputation will be performed as follows (these conventions may change):
  - If only the day of the month is missing, the last day of the month will be used to replace the missing day. If the imputed date is after the death date or the last known alive date, then the latest known alive date or death date is considered as the resolution date.
  - If the day and month are missing or a date is completely missing, it will be considered as missing.

- Missing and partial non-study medication domain dates will be imputed using the derivation algorithm described in 4.1.3 of BMS Non-Study Medication Domain Requirements Specification[17].
- Missing and partial radiotherapy and surgery dates will be imputed using algorithm described in APPENDIX 2.
- For death dates, the following conventions will be used for imputing partial dates:
  - If only the day of the month is missing, the 1st of the month will be used to replace the missing day. The imputed date will be compared to the last known alive date and the maximum will be considered as the death date.
  - If the month or the year is missing, the death date will be imputed as the last known alive date.
  - If the date is completely missing but the reason for death is present, the death date will be imputed as the last known date alive.
- For date of progression after start of study therapy, the following conventions will be used for imputing partial dates:
  - If only the day of the month is missing, the 1st of the month will be used to replace the missing day. In case of the date of death is present and complete, the imputed progression date will be compared to the date of death. The minimum of the imputed progression date and date of death will be considered as the date of progression.
  - If the day and month are missing or a date is completely missing, it will be considered as missing.
- For date of progression to prior therapies, the following conventions will be used for imputing partial dates:
  - If only the day of the month is missing, the 1st of the month will be used to replace the missing day.
  - If the day and month are missing or a date is completely missing, it will be considered as missing.
- For other partial/missing dates, the following conventions were used:
  - If only the day of the month is missing, the 15th of the month will be used to replace the missing day.
  - If both the day and the month are missing, “July 1” will be used to replace the missing information.
  - If a date is completely missing, it will be considered as missing.

The following conversion factors will be used to convert days to months or years:

1 month = 30.4375 days and 1 year = 365.25 days.

Duration (e.g. time-to onset, time-to resolution) will be calculated as follows:

Duration = (Last date - first date + 1)

Last known alive date will be defined based on all appropriate dates collected on the CRF.

All statistical analyses will be carried out using SAS (Statistical Analysis System software, SAS Institute, North Carolina, USA) unless otherwise noted.

9 CONTENT OF REPORTS

All analyses described in this SAP will be included in the Clinical Study Report(s) except where otherwise noted. Refer to the Data Presentation Plan for mock-ups of all tables and listings.

10 DOCUMENT HISTORY

Table 10-1: Document History

Version Number	Author(s)	Description
1	Burcin Simsek	Original issue
2	Burcin Simsek	Employ hierarchical strategy to ORR after PFS and OS. Biomarker and QoL sections are updated slightly. Wording of cabozantinib dose delay and dose reductions are cleaned.

APPENDIX 1 TIME-TO ONSET AND TIME-TO RESOLUTION DEFINITION AND CONVENTIONS FOR SELECT ADVERSE EVENTS, IMMUNE-MEDIATED ADVERSE EVENTS AND EVENTS OF SPECIAL INTEREST

Time-to onset definition

Time-to onset of AE (any grade) for a specific category is defined as the time between the day of the first dose of study treatment and the onset date of the earliest AE (of any grade) in this category.

The time-to onset of AE (grade 3-5) for a specific category is defined similarly with an onset date corresponding to a grade 3-5 AE.

Time-to onset of drug-related AE (any grade or grade 3-5) for a specific category is defined similarly but restricted to drug-related AE.

Time-to onset for a specific subcategory is defined similarly but restricted to event of this subcategory.

Time-to resolution definition

In order to derive the time-to resolution, overlapping or contiguous AEs within a specific category or subcategory will be collapsed into what will be termed "clustered" AEs. For example, if a subject (without pre-treatment AE) experienced an AE from 1st to 5th January, another AE (with different PT but within same category) from 6th to 11th January and same AE from 10th to 12th January, these will be collapsed into one clustered AE from 1st to 12th January.

Key derivation steps for each type of clustered AEs is summarized as follows:

- For any grade AE: Collapse any on-treatment AE from the same category

- For drug-related any grade AE: Collapse any on-treatment drug-related AE from the same category
- For grade 3-5 AE: Collapse any on-treatment AE from the same category. Resolution will be based on the onset date of the earliest grade 3-5 records (if no grade 3-5 record, clustered AE is excluded)
- For drug-related grade 3-5 AE: Collapse any on-treatment drug-related AE from the same category. Resolution will be based on the onset date of the earliest grade 3-5 records (if no grade 3-5 record, clustered AE is excluded)

Time-to resolution of AE (any grade) for a specific category is defined as the longest time from onset to complete resolution or improvement to the grade at baseline among all clustered AEs experienced by the subject in this category per adverse event criteria category. Events which worsened into grade 5 events (death) or have a resolution date equal to the date of death are considered unresolved. If a clustered AE is considered as unresolved, the resolution date will be censored to the last known alive date. Improvement to the grade at baseline implies that all different events in the clustered adverse event should at least have improved to the corresponding (i.e. with same preferred term) baseline grade. This measure is defined only for subjects who experienced at least one AE in the specific category.

The time-to resolution of AE (grade 3-5) for a specific category is defined similarly with an onset date corresponding to a grade 3-5 AE.

Time-to resolution of drug-related AE (any grade or grade 3-5) for a specific category is defined similarly but restricted to drug-related AE.

The time-to resolution of AE (any grade or grade 3-5, drug-related or all) where immune modulating medication was initiated is defined similarly. For data presentation not restricted to IMAE, the additional condition that the subject started an immune modulating medication during the longest AE resolution period will be applied.

Time-to resolution for a specific subcategory is defined similarly but restricted to event of this subcategory.

The algorithm for collapsing adverse event records is using the following conventions:

For each subject and specified category, the corresponding adverse event records will be collapsed when:

1) Multiple adverse event records have the same onset date.
2) The onset date of an event record is either the same day or 1 day later than the resolution date of a preceding event record (contiguous events).
3) The onset date of an event record is after the onset date and prior to or on the resolution date of a preceding event record (overlapping events).

APPENDIX 2 MISSING AND PARTIAL RADIOTHERAPY AND SURGERY DATES IMPUTATION ALGORITHMS

Procedures – Imputation Rules.

If reported procedure start date is a full valid date then set start date equal to the date part of procedure start date.

In case of partial date use imputation rules described below:

- If only day is missing then
  - If month and year of procedure match month and year of first dose date then impute as date of first dose;
  - If month and year of procedure don't match month and year of first dose date then impute as first day of that month and year.
- If both day and month are missing, then impute as maximum between 01JAN of the year and date of the first dose;
- If date is completely missing or invalid then leave missing.

Note: Imputation is not applicable to data where start date is not collected (for example "PRIOR RADIOTHERAPY" CRF). Set start date to missing in this case.

If reported end date is a full valid date then set end date equal to the date part of the reported end date.

In case of partial date use imputation rules described below:

- If reported end date is partial then set end date equal to the last possible reported end date based on the partial entered reported end date.
- If reported end date is missing, continuing, unknown or invalid then set end date equal to the most recent database extraction date.

If end date was imputed then compare end date to the death date or last known alive date if subject is not dead. If posterior then end date should be imputed to death date (or last known alive date if subject not dead).

Note: Imputation of partial dates only applies to data entered on "RADIOTHERAPY" CRF page. For other CRF pages in case of partial dates set end date to missing.

Surgeries – Imputation Rules.

If reported surgery date is a full valid date then set start date equal to the date part of surgery date.

In case of partial date, use one of the two imputation rules described below:

A. For data collected on "PRIOR SURGERY RELATED TO CANCER" CRF page:

- If only day is missing then impute as the first day of the month;
- If both day and month are missing then then impute as 01JAN of the year;

- If date is completely missing or invalid then leave missing.

B. For data collected on other CRF pages (deemed to be on-treatment/subsequent surgeries):

- If only day is missing then
  - If month and year of surgery match month and year of first dose date then impute the missing date as the date of first dose;
  - If month and year of surgery don't match month and year of first dose date then impute as first day of that month and year;
- If both day and month are missing then impute as maximum between 01JAN of the year and date of the first dose;
- If date is completely missing or invalid then leave missing.

APPENDIX 3 IMMUNOGENICITY ANALYSIS: BACKGROUND AND RATIONALE

The following summary is from the FDA Guidance for Industry Immunogenicity Assessment for Therapeutic Protein Products and White Paper on Assessment and Reporting of the Clinical Immunogenicity of Therapeutic Proteins and Peptides – Harmonized Terminology and Tactical Recommendations by Shankar et al. The program-level definitions of sample- and subject-level ADA status are based on recommendation from the BMS Immunogenicity Council.

Immune responses to therapeutic protein products may pose problems for both subject safety and product efficacy. Immunologically based adverse events, such as anaphylaxis and infusion reactions, have caused termination of the development of therapeutic protein products or limited the use of otherwise effective therapies. Unwanted immune responses to therapeutic proteins may also neutralize the biological activity of therapeutic proteins and may result in adverse events not only by inhibiting the efficacy of the therapeutic protein product, but by cross-reacting to an endogenous protein counterpart, if present. Because most of the adverse effects resulting from elicitation of an immune response to a therapeutic protein product appear to be mediated by humoral mechanisms, circulating antibody has been the chief criterion for defining an immune response to this class of products.

ADA is defined as biologic drug-reactive antibody, including pre-existing host antibodies that are cross-reactive with the administered biologic drug (baseline ADA). Titer is a quasiquantitative expression of the level of ADA in a sample. By employing a serial dilution-based test method, titer is defined as the reciprocal of the highest dilution of the sample (e.g., dilution of 1/100 = titer of 100). The ADA is also tested, via a cell-based biologic assay or a non cell-based competitive ligand-binding assay for a subpopulation of ADA known as neutralizing antibodies (NAb), which inhibits or reduces the pharmacological activity of the biologic drug molecule regardless of its in vivo clinical relevance. Non-neutralizing ADA (non-NAb) is ADA that binds to the biologic drug molecule but does not inhibits its pharmacological activity.

ADA should be tested using sensitive and valid methods and employing an appropriate strategy for elucidating immunogenicity. Detection of ADA is typically performed in three tiers (screening, confirmatory, and titer) using statistically determined cutpoints and samples testing positive in the ADA assay are analyzed for neutralizing activity, especially in late-stage clinical studies. “Detection” of ADA implies that drug-specific ADA was confirmed. The ‘drug tolerance” of an assay (highest drug concentration that does not interfere in the ADA detection method) is not an absolute value and differs between individuals due to the varying avidities of ADA immune responses. An ADA sampling strategy of collecting samples at times when the least drug concentration is anticipated (trough concentrations) can increase the likelihood of accurate ADA detection.

It is useful to present ADA results from clinical studies as (a) characteristics of the ADA immune response, (b) relationship of ADA with pharmacokinetics (PK) and, when relevant, pharmacodynamics (PD) biomarkers, and (c) relationship of ADA with clinical safety and efficacy.

Clinical consequences of ADA can range from no apparent clinical effect to lack of efficacy (primary treatment failure), loss of efficacy (secondary treatment failure) or heightened effect due to altered exposure to the biologic drug, adverse drug reactions (administration-related systemic or site reactions), and severe adverse drug reactions (anaphylaxis and unique clinical problems associated with cross-reactivity and neutralization of endogenous molecules). Thus it becomes important to examine any associations between ADA or any of its attributes with the various clinical sequelae. The presence of ADA may or may not preclude the administration of drug to ADA-positive subjects because the outcome is dependent upon the magnitude of the impact of ADA on PK and PD. Hence, the relationship of ADA with PK/PD is an important additional consideration, but does not necessarily result in a clinically impactful consequence per se.

Immunogenicity Endpoints

A fundamental metric that informs clinical immunogenicity interpretation is the incidence of ADA in a study or across comparable studies. ADA incidence is defined as the proportion of the study population found to have seroconverted or boosted their pre-existing ADA during the study period.

Terms and Definitions

Validated ADA test methods enable characterization of samples into ADA-positive vs. ADA-negative. To classify the ADA status of a subject using data from an in vitro test method, each sample from the subject is categorized based on the following definitions:

Sample ADA Status:

- Baseline ADA-positive sample: ADA is detected in the last sample before initiation of treatment
- Baseline ADA-negative sample: ADA is not detected in the last sample before initiation of treatment
- ADA-positive sample: After initiation of treatment, (1) an ADA detected (positive seroconversion) sample in a subject for whom ADA is not detected at baseline, or (2) an ADA detected sample with ADA titer to be at least 4-fold or greater (≥) than baseline positive titer
- ADA-negative sample: After initiation of treatment, ADA not positive sample relative to baseline

Next, using the sample ADA status, subject ADA status is defined as follows:

Subject ADA Status:

- Baseline ADA-positive subject: A subject with baseline ADA-positive sample
- **ADA-positive subject**: A subject with at least one ADA positive-sample relative to baseline at any time after initiation of treatment
  1) *Persistent Positive (PP):* ADA-positive sample at 2 or more consecutive time points, where the first and last ADA-positive samples are at least 16 weeks apart
  2) *Not PP-Last Sample Positive:* Not persistent positive with ADA-positive sample at the last sampling time point

11) *Other Positive:* Not persistent positive but some ADA-positive samples with the last sample being negative
12) *Neutralizing Positive:* At least one ADA-positive sample with neutralizing antibodies detected

- **ADA-negative subject:** A subject with no ADA-positive sample after the initiation of treatment.

(Note: 16 weeks was chosen based on a long half-life of IgG4.)

Population for Analyses

Analysis of immunogenicity data will be based on ADA evaluable subjects defined as all treated subjects with baseline and at least 1 post-baseline immunogenicity assessment. Analysis dataset and data listing will include all available ADA samples. However, subject-level ADA status will be defined based on only adequate samples (e.g., excluding 1-hour post-infusion samples when clearly indicated).

11 REFERENCES

[1] Apolo AB, Tomita Y, Lee MJ, Lee S, Frosch A, Steinberg SM. Effect of Cabozantinib on immunosuppressive subsets in metastatic urotherlial carcinoma [abstract No. 4501]. 2014. J Clin Oncol. 2014; 32 (5s). 2014 ASCO Annual Meeting.

[2] Dolan P. Modeling valuations for EuroQol health states. Medical Care 1997;35: 1095-1108

[3] The EuroQol Group: EuroQol: A new facility for the measurement of health-related quality of life—The EuroQol Group. Health Policy 16:199-208, 1990.

[4] Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. Health Qual Life Outcomes 2007;5:70.

[5] Brookmeyer R. and Crowley J. A confidence interval for the median survival time. Biometrics 38:29-41, 1982

[6] Klein, J. P. and Moeschberger, M. L. (1997), Survival Analysis: Techniques for Censored and Truncated Data, New York: Springer-Verlag

[7] Clopper, CJ and Pearson, ES. The use of confidence or fiducial limits illustrated in the case of the binomial.Biometrika 26: 404-423, 1934.

[8] Fagerland MW, Lydersen S, Laake P. Recommended confidence intervals for two independent binomial proportions. Stat Methods Med Res 2011. Published Online First 13 Oct 2011.

[9] Global Biometric Sciences, SAS Analysis Dataset Specification, Unique Adverse Events. Version 2.0, April 22, 2015

[10] Glimm E, Maurer W, Bretz F. Hierarchical testing of multiple endpoints in group-sequential trials. Statistics in Medicine. 2010;29:219-228

[11] Greenwood, M. The errors of sampling of the survivorship tables, Reports on Public Health and Statistical Subjects, 33, Appendix 1, HMSO, London, 1926

[12] Kalbfleisch, J. D. and Prentice, R. L. (1980), The Statistical Analysis of Failure Time Data, New York: John Wiley & Sons.

[13] Fleming TR, Harrington DP 1991: Counting processes and survival analysis. New York: John Wiley and Sons

[14] Gail, M. and Simon, R. “Testing for qualitative interactions between treatment effects and subject subsets”. Biometrics, 1985; 41:361-372.

[15] Clopper, C.; Pearson, E. S. (1934). "The use of confidence or fiducial limits illustrated in the case of the binomial". *Biometrika* **26**: 404–413.

[16] Adverse Event Domain Requirements Specification Bristol Myers Squibb Co. PRI. Version 2.3.0 April 23, 2018

[17] Non-Study Medication Domain Requirements Specification Bristol Myers Squibb Co. PRI. Version 2.10.0 April 23, 2018